TRAC-M-TM-17-021
April 2017

Cognitive Agility Measurement in a Complex Environment

TRADOC Analysis Center-Monterey
700 Dyer Road
Monterey, CA 93943-0692

This study cost the
Department of Defense approximately
$183,000 expended by TRAC in
Fiscal Years 15-17.
Prepared on 20170411
TRAC Project Code # 060315

This page left intentionally blank.

TRAC-M-TM-17-021
April 2017

Cognitive Agility Measurement in a Complex Environment

Authors

MAJ Adam Haupt
Dr. Quinn Kennedy
Dr. Samuel Buttrey
Dr. Jonathan Alt
LT. Miriam Mariscal
LT. Lindayle Fredrick

TRADOC Analysis Center-Monterey
700 Dyer Road
Monterey, CA 93943-0692

This page left intentionally blank.

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.			
1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** 11-04-2017	**3. REPORT TYPE AND DATES COVERED** Technical Memorandum, August 2015 – April 2017	
4. TITLE AND SUBTITLE Cognitive Agility Measurement in a Complex Environment			**5. PROJECT NUMBERS** TRAC Project Code 060315
6. AUTHOR(S) MAJ Adam Haupt, Dr. Quinn Kennedy, Dr. Samuel Buttrey, Dr. Jonathan Alt, LT. Miriam Mariscal, LT. Lindayle Fredrick			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) US Army TRADOC Analysis Center - Monterey 700 Dyer Road Monterey CA, 93943-0692			**8. PERFORMING ORGANIZATION REPORT NUMBER** TRAC-M-TM-17-021
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) Headquarters, TRAC ATTN: ATRC 255 Sedgwick Avenue Fort Leavenworth, KS 66027-2345			**10. SPONSORING/MONITORING AGENCY REPORT NUMBER** TRAC-M-TM-17-021
11. SUPPLEMENTARY NOTES Findings of this report are not to be construed as an official Department of the Army (DA) position unless so designated by other authorized documents.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT DISTRIBUTION STATEMENT A: Approved for public release; distribution is unlimited.			**12b. DISTRIBUTION CODE**
13. ABSTRACT (maximum 200 words) The purpose of this project is to operationalize and measure the Cognitive Agility levels of military decision makers in order to support the Army's line-of-effort (LOE) of Cognitive Dominance. The research team creates a human cognitive experiment using psychological tests and a military decision computer game called Make Goal to attempt to measure cognitive agility in military leaders. The experiment is conducted on 40 Naval Postgraduate School (NPS) students. Performance and eye-tracking data is collected and analyzed by two NPS thesis students. This document discusses the experimental design and the results from one of those theses.			
14. SUBJECT TERMS cognitive agility, human dimension, eye-tracking, psychology, data analysis, military computer game, human experimentation, stress			**15. NUMBER OF PAGES** 20 **16. PRICE CODE**
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

This page left intentionally blank.

NOTICES

DISCLAIMER

Findings of this report are not to be construed as an official Department of the Army (DA) position unless so designated by other authorized documents.

REPRODUCTION

Reproduction of this document, in whole or part, is prohibited except by permission of the Director, TRAC, ATTN: ATRC, 255 Sedgwick Avenue, Fort Leavenworth, Kansas 66027-2345.

DISTRIBUTION STATEMENT

Approved for public release; distribution is unlimited.

DESTRUCTION NOTICE

When this report is no longer needed, DA organizations will destroy it according to procedures given in AR 380-5, DA Information Security Program. All others will return this report to Director, TRAC, ATTN: ATRC, 255 Sedgwick Avenue, Fort Leavenworth, Kansas 66027-2345.

This page left intentionally blank.

ABSTRACT

The purpose of this project is to operationalize and measure the Cognitive Agility levels of military decision makers in order to support the Army's line-of-effort (LOE) of Cognitive Dominance. The research team creates a human cognitive experiment using psychological tests and a military decision computer game called Make Goal to attempt to measure cognitive agility in military leaders. The experiment is conducted on 40 Naval Postgraduate School (NPS) students. Performance and eye-tracking data is collected and analyzed by two NPS thesis students. This document discusses the experimental design and the results from one of those theses.

This page left intentionally blank.

Table of Contents

List of Figures

This page intentionally left blank.

Chapter 1 – Introduction

1.1. Purpose

The purpose of this project is to operationalize and measure the Cognitive Agility levels of military decision makers in order to support the Army's line-of-effort (LOE) of Cognitive Dominance.

1.2. Background

In studying the human dimension aspect of the Army Operating Concept, the Army identified three lines of effort: "establish cognitive dominance, execute realistic training, and drive institutional agility" (U.S.A. CAC, 2014). The focus of the research in this project is on cognitive dominance, defined as "a position of intellectual advantage over a situation or adversary that fosters proactive agility over reactive adaptation, facilitating the ability to anticipate change before it occurs" (U.S.A. CAC, 2014). Currently the Army does not have the capability to adequately measure and evaluate the cognitive agility levels of its leaders. This research attempts to measure it by utilizing statistical and neuropsychological measures of cognitive agility in a military context.

1.3. Problem Statement

The Army cannot scientifically measure and assess the cognitive agility levels of a military decision maker.

1.4. Hypothesis

By using a combination of respected psychological tests and a military relevant computer game the study team will be able to scientifically measure different levels of cognitive agility in military leaders.

1.5. Study Issues and Essential Elements for Analysis

Issue 1: Can the three components of Cognitive Agility be measured in a computer military wargame (Make Goal) using statistically significant metrics?

EEA 1.1: Is the cognitive flexibility independent variable in Make Goal statistically significant in the model?

EEA 1.2: Is the cognitive openness variable in Make Goal statistically significant in the model?

EEA1.3: Is the focused attention variable in Make Goal statistically significant in the model?

Issue 2: Are historical psychological tests statistically significant in the successful outcome of Make Goal?

EEA 2.1: Is the cognitive flexibility variable from the Stroop test statistically significant?

EEA 2.2: Is the cognitive openness variable from the Alternate Use Test (AUT) statistically significant in the model?

EEA2.3: Is the focused attention variable from the Go, No Go Paradigm statistically significant in the model?

Issue 3: Do the measures developed for Make Goal correlate with their corresponding historical psychology tests?

EEA3.1: Does the variable for Make Goal cognitive flexibility correlate with the Stroop Test cognitive flexibility variable?

EEA3.2: Does the variable for Make Goal cognitive openness correlate with the AUT cognitive openness variable?

EEA3.3: Does the variable for Make Goal focused attention correlate with the Go, No Go Paradigm focused attention variable?

1.6. Constraints, Limitations and Assumptions

- Constraints[1].
  - The project timeline is 17 AUGUST 2015 – 26 APRIL 2017.
  - Study method must be approved by Internal Review Board (IRB).
- Limitation[2].
  - Study team must use at least one human-in-the-loop military strategy game.
  - Cognitive flexibility is the primary in-game measurement captured in the Make Goal experiment.
  - Human-in-the-loop testing is limited to active duty Naval Postgraduate School (NPS) students, faculty, and staff and TRAC-MTRY personnel.
- Assumptions[3].
  - Military strategy game can be modified to isolate, at a minimum, one component of cognitive agility.
  - Traditional cognitive agility tests correlate with the results of the Make Goal.
  - The NPS sample will provide an accurate representation of military decision makers.

[1] Constraints limit the study team's options to conduct the study.
[2] Limitations are a study team's inability to investigate issues within the sponsor's bounds.
[3] Assumptions are study-specific statements that are taken as true in the absence of facts.

Chapter 2 – Methodology, Analysis and Implementation

2.1. Methodology Overview

In this chapter we examine the methodology of this project. Our methodology includes six steps (see Figure 1). We execute this project in two phases. In the phase one the project team conductes steps 1-3 in order to scope the project and build the necessary cognitive agility collection tools needed in the experiment. In phase two the study team uses two NPS thesis students to run the tests on 40 NPS student volenteers, collect data and do the final analysis.

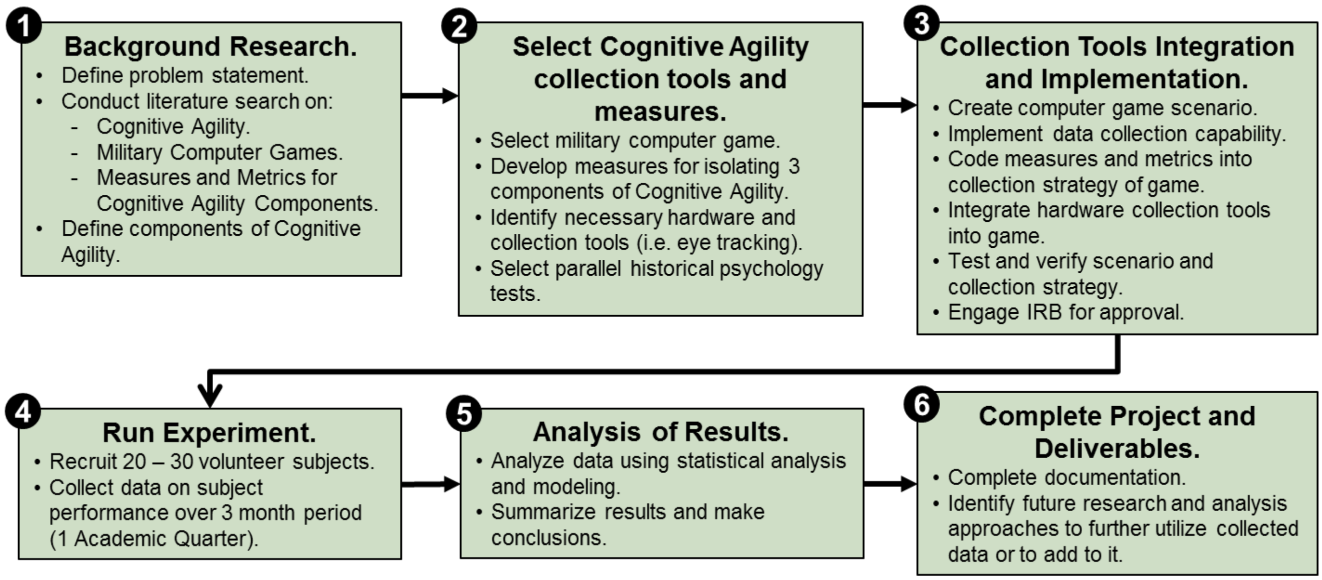


Figure 1. Methodology for Cognitive Agility project.

2.2. Background Research

In Dr. Darren Good's paper "Explorations of Cognitive Agility: Real time Adaptive Capacity," he defines cognitive agility to be a construct with three components that are:

- Cognitive Flexibility is the ability to cognitively control and shift mental set and overcome dominant or automatic response sets.

- Cognitive Openness is being receptive to new ideas, experiences and perspectives. Individuals with this trait will be more prone to exploring new creative solutions.

- Focused Attention describes the ability to attend to relevant stimuli and ignore distracting ones.

In his research he provides a methodology for measuring adaptive learning using the components of cognitive agility and measuring them with psychological tests and an interactive computer game. Among his tests he uses the Stroop test to measure cognitive flexibility, Go/No Go test to measure focused attention, and the AUT to measure cognitive openness. These tests proved significant in measuring those components and are correlated to adaptive performance metrics in his interactive computer game.

The idea that cognitive function can be tested and measured with computer games is further supported in the paper “Supporting cognitive adaptability through game design” (Gallager, P.S., & Prestwich, S.H., 2012). This paper defines game design features that are recommended for measuring cognitive adaptability. The lessons learned from this research draw heavily from the Wisconsin Card Sorting Test (WCST), which is game where the subject attempts to match a card to a choice of four other cards. The subject does not know the rules by which he is supposed to match the cards, but is given responses of correct or incorrect based on his selection. As the subject begins to learn the card matching rules the rules are periodically changed and the subject must identify that the conditions are changed and discover the new rules. This research states that in order to measure cognitive adaptability the game must have certain key features that are demonstrated in the WCST. They are unstated rules, dynamic shifting environments, open-ended game play and implicit reinforcement for individual actions or choices to achieve the final goal.

A demonstration of how a psychological, interactive computer game can be repurposed for military decision makers is found in the paper “Iowa Gambling Task Modified for Military Domain”, (Nesbit, Kennedy, Alt, Fricker, 2015). This research states that reinforced learning, which is critical for optimal decision making, can be measured in military decision makers by repurposing the Iowa Gambling Task (IGT). In the IGT, participants receive a loan of $2,000 of play money and are asked to make a series of decisions to maximize the profit on the loan. Each decision entails selecting one card at a time from any of four available decks of cards. All cards give money and some cards also issue a penalty. Decks differ in the amount of money given on a single trial ($50 or $100) as well as the frequency and severity of penalties ($0 to $1,250). Healthy participants should learn through reinforcement learning which decks have the best long term payoffs (Bechara et al., 1994; Steingroever et al., 2013). In this study the IGT is modified for military use by replacing the cards with pictures of identical convoy routes. Feedback is given by rewarding the player with a value of damage to the enemy forces for choosing a good route or penalizing the player with damage to friendly forces for choosing a bad route.

Although many other documents address cognitive function and ways to measure it, the above research was the most influential in the study design of this project. Dr. Good’s research helped the study team decide on how to define and measure cognitive agility. The other two research efforts helped the study team to combine the principals of the WCST and the IGT and modify them to fit the military domain.

2.3. Experimental Design

The experimental design consists of six tests meant to measure cognitive agility (see figure 2). There are two main categories of tests. The first category consists of validated psychological tests that have been historically used to measure the components of cognitive agility. The second category consists of tests that are also designed to measure cognitive agility but are tied to a military computer game named Make Goal.

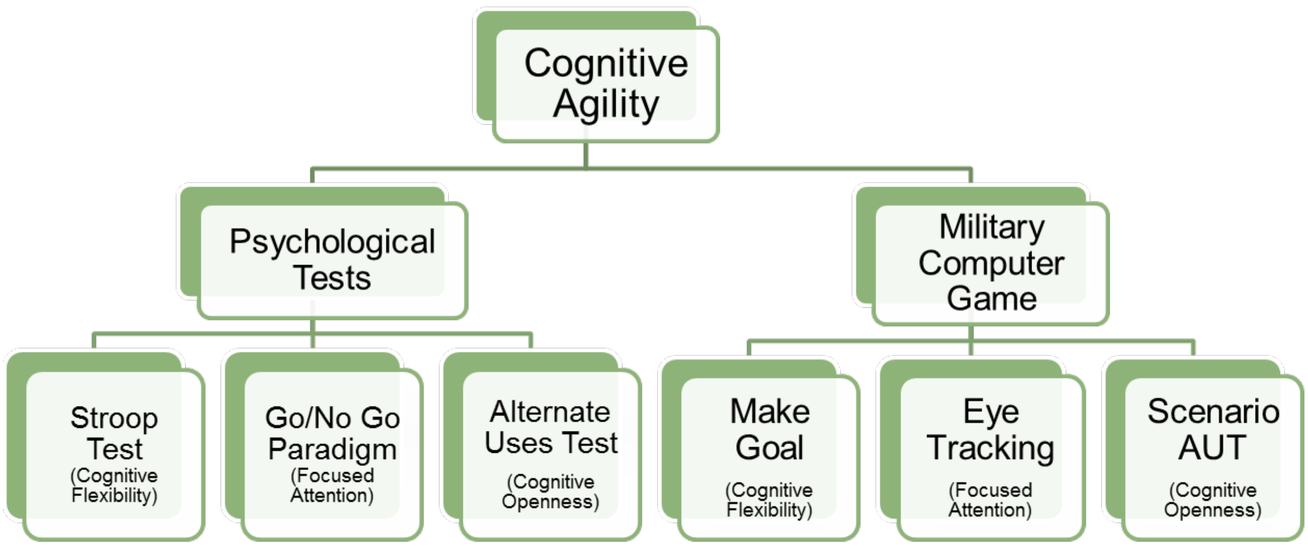


Figure 2. Cognitive Agility Test Architecture.

2.3.1. Stroop Test

The Stroop Task (Stroop, 1935) measures cognitive flexibility as a performance score. The word color Stroop task challenges the participant to respond to a font color in which an incongruent word of a color is presented (e.g., the word blue written in red font). The participant is instructed to select the word of the font color (red) instead of the written word itself (blue), (see figure 3). As responding to the word spelling rather than font color is the automatic reaction (MacLeod, 1991), it serves as a measure of the participant's ability to override and flexibly respond.

We develop this test in the software package PsychoPy™ using four colors (red, blue, green, orange) in 16 font color / word spelling combinations. The participant is presented 80 font/word challenges and the response time and number of correct answers is documented. Metrics from this test include total number or correct responses and the average response time.

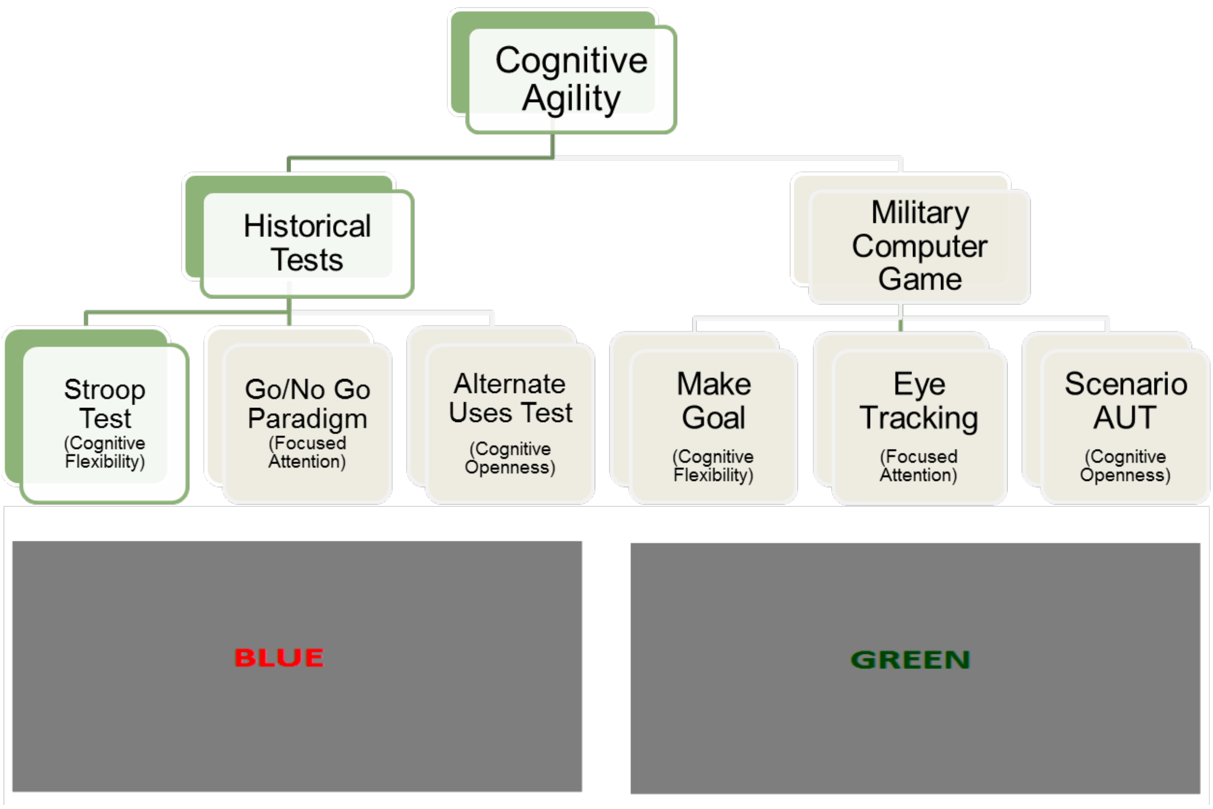


Figure 3. Stroop Test visualization example.

2.3.2. Go/No Go Paradigm

The go/no go paradigm measures focused attention (Zimmermann & Fimm, 2000). Participants are challenged to respond (by pressing the space bar) as fast as possible when a 'go' stimuli appears and to refrain from responding (by not pressing the space bar) when 'no go' stimuli appear. Participants are asked to memorize two 3 × 3 textured squares (go stimuli). Then squares appear that are the same (go stimuli) and slightly different (no go stimuli). Reaction time (at the level of milliseconds) is used as the performance score. Reliability for go/no go has been demonstrated with split half and odd even coefficients at 0.998 (Zimmerman & Fimm, 2000). Performance scores on the go/no go correlate with the Barrett Impulsiveness Scale ($r = 0.40$, $p<.01$) and perseverative error on the Wisconsin Card Sorting Task ($r = -0.46$) (Keilp, Sackeim, & Mann, 2005).

This test is also developed in PsychoPy™. The 'go' stimulus is a green vertical or horizontal rectangle and the 'no' stimulus is a blue vertical or horizontal rectangle, see figure 4. The participant has a varying amount of time to respond (between .1 and .5 seconds). Metrics from this test are total number of correct answers and response time.

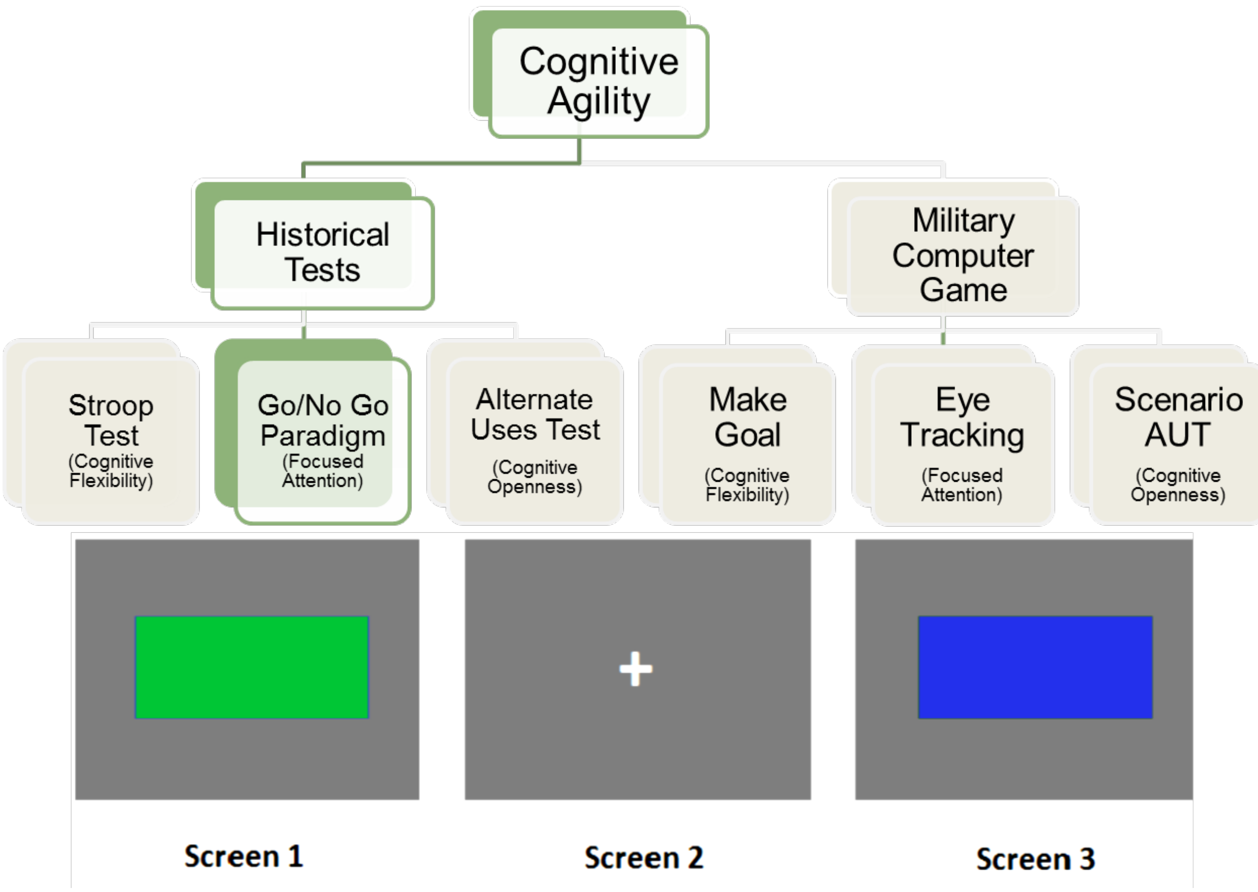
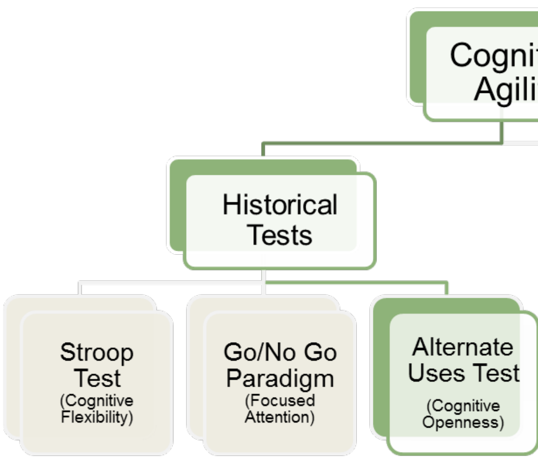


Figure 4. Go/No Go visualization example.

2.3.3. Alternate Use Test (AUT)

The Alternate Uses Test (AUT) challenges the participant to list as many possible uses for a common item in a timed setting. Our test used a brick for the common item and participants are given four minutes to list all the creative uses that they can think of for that item, see figure 5. The number of items generated is used as the cognitive openness performance score. The AUT correlates significantly with openness on the NEO (r = 0.46; Chamorro- Premuzic, 2006), the Barron Symbolic Equivalence Test (r = 0.49; Barron, 1988) and with greater sensitivity in a habituation process (r = 0.36; Martindale, Anderson, Moore, & West, 1996). The metric that we gather from this test is the total number of feasible/creative answers. Other metrics can be gathered from this test such as a score for originality (how unusual the uses are), flexibility (range of ideas, in different domains) and elaboration (level of detail and development of the idea), but these require more subjective evaluation.



Example Answers to AUT Test	
1.	build a house
2.	weapon
3.	hammer
4.	break it up into small pieces to form a sharp edge for cutting.
5.	break it up even further and grind it down to particle size to use as putty to seal a crack.
6.	line a sidewalk
7.	wall
8.	build a fountain

Figure 5. AUT example responses for a brick.

2.3.4. Make Goal

Make Goal was originally developed by the MOVES Institute at NPS in support of Navy Recruiting Command. It challenged recruiting commanders to understand their regional demographics, manage their recruiters and make their recruiting goals. The number monthly accessions were determined by an accessions function based off of historical data. For this research we modify Make Goal so that the game is a cross between IGT and WCST. It is similar to IGT in that a player must select one of four choices each turn in an attempt to get the highest score possible. It also shares the rule changing aspect of the WCST that tests for cognitive flexibility.

In the game players have 100 turns to recruit as many recruits as possible. The player achieves this by placing a recruiter each turn in one of four regions in Texas. Depending on their selection for that turn they recruit some number of recruits. The number of recruits for each region is determined by four different distributions. The challenging aspect of the game is determining through trial and error which regions yield the best results. This aspect of the game is similar to the IGT in that a player must learn through repetition which choices yield the optimal results that are also based on a repeating distribution of numbers. This task is made more difficult by systematic rule changes where the distributions are switched to other regions. At these times a player must identify the rule change and begin searching for the best regions. This task requires cognitive flexibility and shares characteristics with the WCST, (see Figure 6 for visual of Make Goal playing screen).

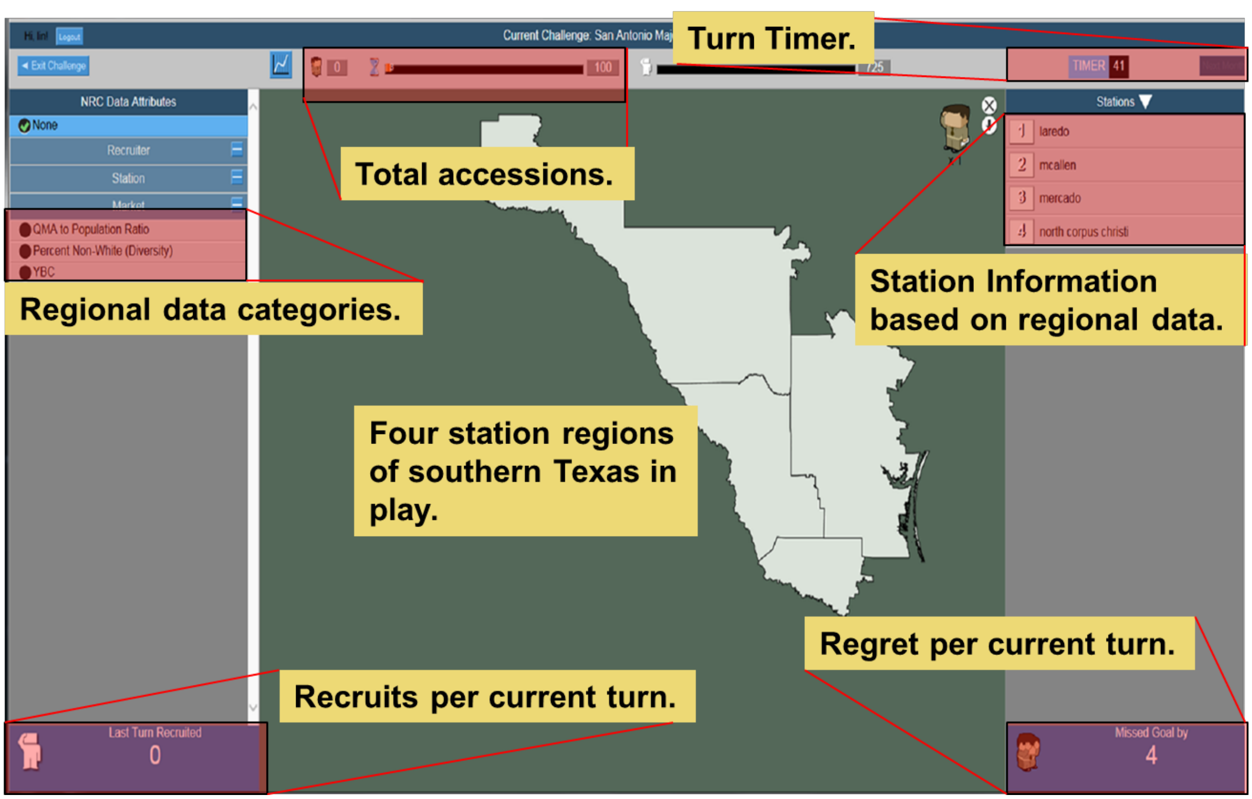


Figure 6. Make Goal game interface.

2.3.4.1. Make Goal Design

Make Goal consists of 100 turns. These turns are broken up into six sets that are 20, 18, 17, 16, 16, 13 turns respectively. The player can place a recruiter on any of four Texas regions each turn. The four regions are Laredo, McAllen, Mercado, and North Corpus. There are four different outcome distributions. One of the outcome distributions is clearly the best on average with an average reward of six accessions. The next best outcome distribution rewards an average of four accessions. The last two distributions are distinctly worse with average accessions of two and one respectively. Which region each distribution is associated with changes each set, (see Figure 7). This is meant to force the player to identify that the conditions of the game have been changed and to take appropriate action to find the best outcome distribution again. This is made more difficult by two factors. First, the conditions change at a quicker rate (fewer turns) as the game progresses. Second, the player has a time limit for each turn that becomes progressively shorter as the game proceeds (100 to 15 seconds). The player is given feedback in three forms. First, the player is shown his level of "reward" in the form of the number of recruits he attained in the last turn, (see Figure 6 bottom left corner). Second, the player is shown some level of "penalty", which is his regret (i.e. the amount of recruits he could have gotten if he had selected the optimal selection for that turn), (shown in the bottom right corner of Figure 6). Third, the player is shown his total number of accessions that he has recruited in the game up to that point, (see top center box in Figure 6). Using this feedback the player is able to iteratively learn the game and find the optimal recruiting region for that set.

Another game design feature is added into the game may aid or possibly confuse the player. Regional data is provided about each of the regions. There are three categories:

- Diversity: Percent of population that was non-white.
- QMA to Population Ratio: Ratio of Qualified Military Aged personnel in the regional population.
- YBC: Number of youth between 17 and 22 years of age.

Each region scores differently in each category and for each set one of these categories is chosen to be the decisive environmental factor that drives recruiting. The distributions are then linked to the regions that best represent the driving category for that set (e.g. in set 1 diversity is the driving recruiting factor. Laredo has the highest diversity so it is linked to the optimal outcome distribution whereas North Corpus is the least diverse and therefore is linked to the least optimal outcome distribution), (see Figure 7).

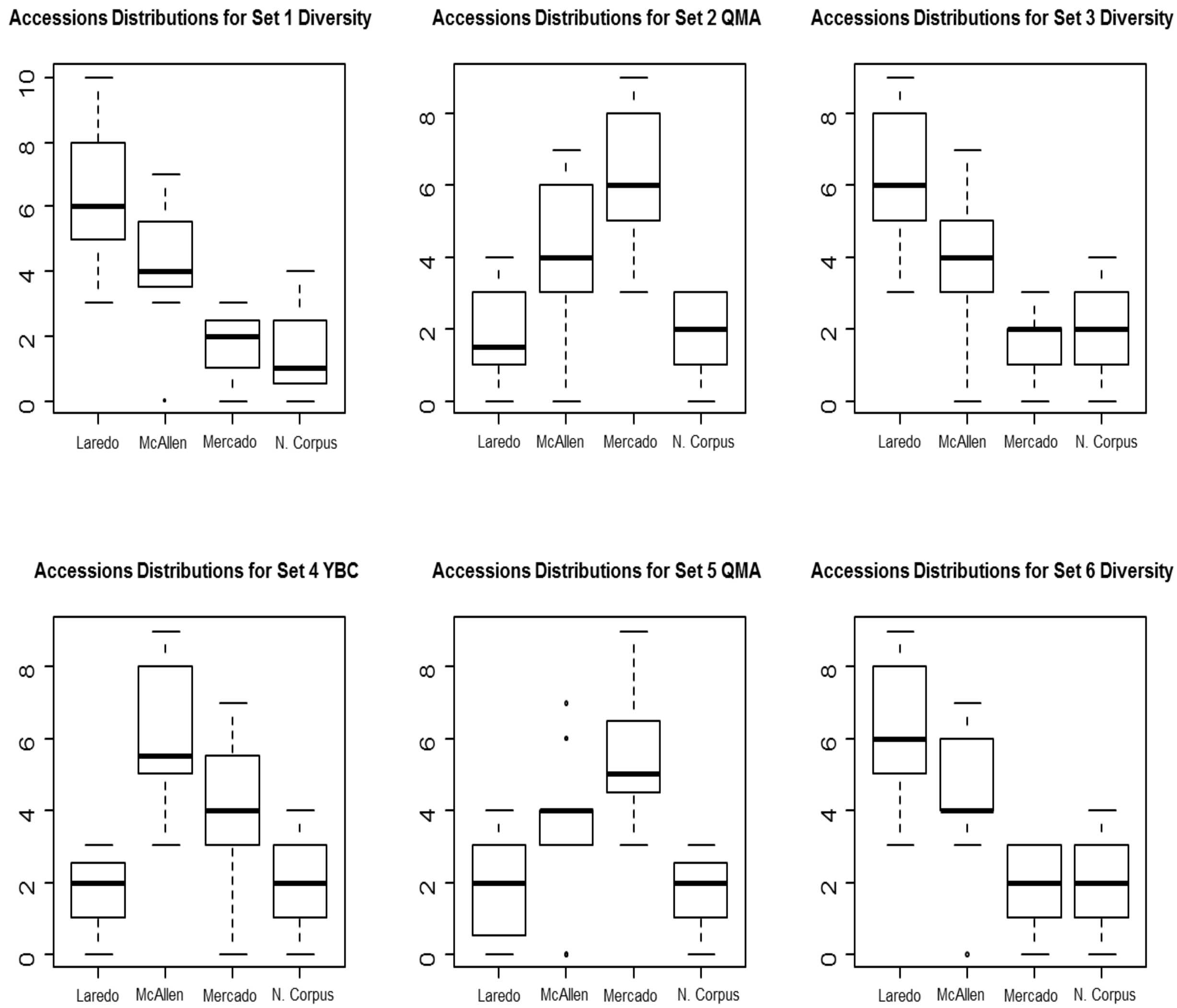


Figure 7. Boxplots of outcome distributions by set and recruiting region.

2.3.4.2. Make Goal Eye-Tracking Design

In addition to collecting performance data on players as they make choices in the Make Goal game, eye-tracking data is also collected to see if different players focus on different game screen regions while playing. This is important to the study because the study team is interested in whether or not the cognitive agility component of focused attention can be measured through a player's eye-tracking patterns. In order to collect this data the game screen is divided into six regions. These regions are shown in Figure 8. During the game the amount of time the player spends focused on each region is recorded so that it can be analyzed against the player's performance. For more information on eye-tracking analysis, see Appendix C.

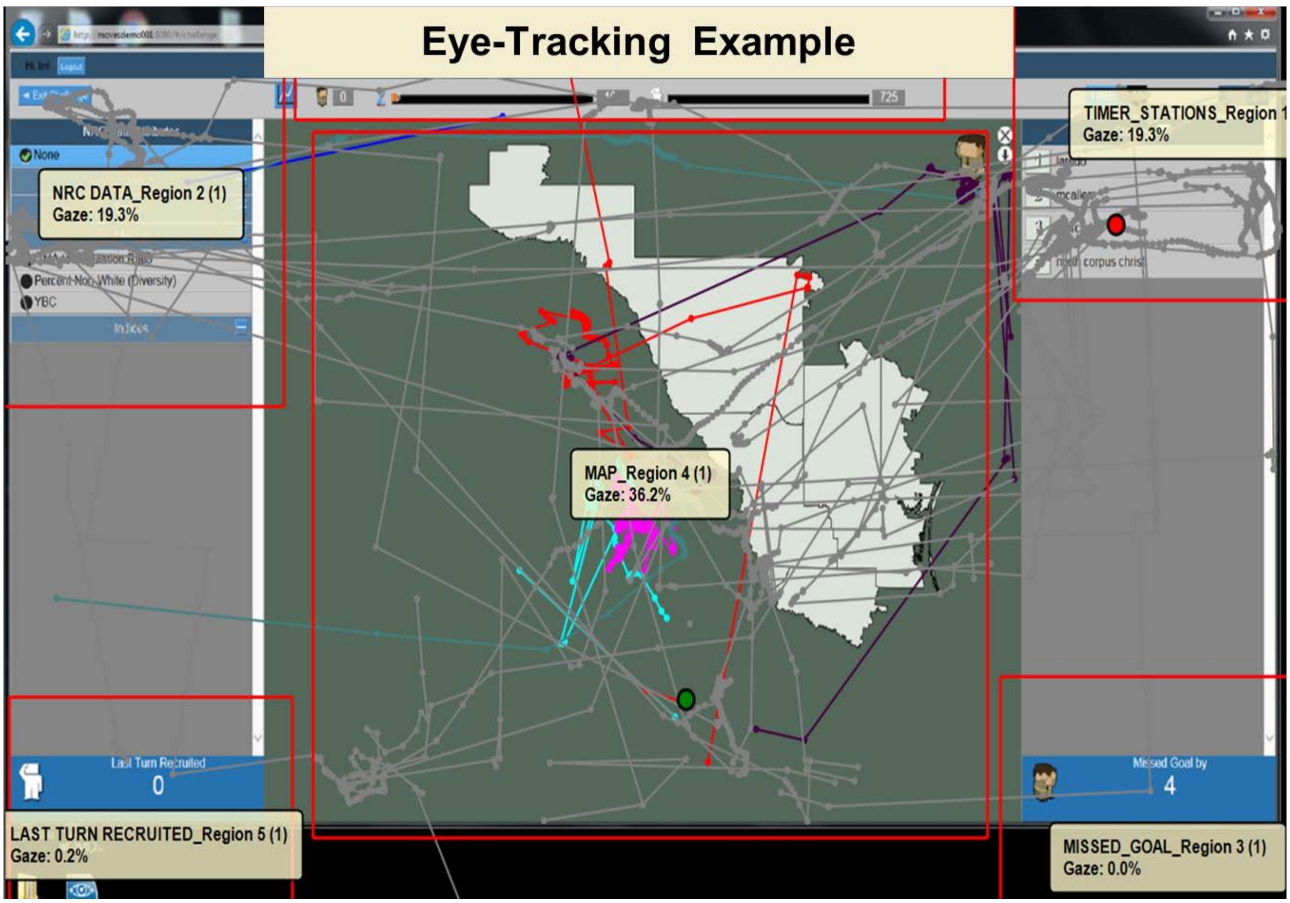


Figure 8. Eye-Tracking regions on the Make Goal game screen.

2.3.4.3. Make Goal Alternate Use Test

Cognitive openness is captured within Make Goal by giving players a modified AUT after completing the Make Goal game. The test is modified by asking the player to list as many factors that they believed influenced the turn-by-turn accessions outcomes, rather than listing as many uses for a common object. The metric returned is the total number of factors listed.

2.3.4.4. Make Goal Post Test Survey

At the end of the test subjects are given a post-test survey consisting of eight questions. This survey asks two types of questions. The first type is open ended questions about the subject's strategy and perceptions of the Make Goal game. The second type is questions about the level of stress and challenge the subject experienced during the experiment. These questions are responded to on a nine point Likert scale. The purpose of this survey is to allow subjects to self-report on the inner workings of their thought process and to see if the stress group did in fact experience more stress than the control group.

2.3.5. Stress and Control Groups

To see if stress influences performance in the Make Goal game, participants are divided into two groups. As participants are selected, they are randomly assigned to a Stress or Control group. The randomly assigned members of the stress group play Make Goal with the experimenter observing them closely and making comments as they play the game. A script is followed by the experimenter to ensure consistency for all subjects. At the following intervals the following comments are made:

- Move 35: “Are you sure you want to make that move?”
- Move 50: “You might want to take a look at the timer.”
- Move 70: “Usually people are doing a little better by this point.”
- Move 85: “Might want to check on that timer again.”
- Move 90: “It doesn't look like you're doing too well.”

The randomly assigned members of the control group play Make Goal without the experimenter observing them closely or making any comments on their performance (Mariscal, 2017).

2.4. Experiment Implementation

The experimental design for this project combines both historically proven psychological tests with our own military decision making game. Once the study team completed the design of the experiment and the collection tools, two thesis students from the Naval Postgraduate Students (NPS) were selected to recruit 40 test subjects from their cohort and collect the data over the course of two months. The subjects were all volunteers and represented military officers from the Army, Navy and Marine Corps (see Table 1 for demographics of the volunteers).

Table 1. Demographics of the 40 experiment participants.

	Control Group	**Stress Group**
Service	USA: 4 USMC: 6 USN:10	USA: 4 USMC: 3 USN: 12 International: 1
Rank	O5: 1 O4: 9 O3: 9 O1: 1	O4: 7 O3: 9 O2: 4
Years in Service (mean / sd)	11.05 / 6.14	9 / 4.76
Age (mean / sd)	33.7 / 6.28	30.7 / 4.49
Gender	14 male 5 female	10 male 10 female

Once the data was collected the analysis for this project was scoped into two separate theses. The first thesis by Miriam Mariscal (Lieutenant, USN) is titled "An Exploration of Cognitive Agility as Quantified by Attention Allocation in a Complex Environment" and aims to 1) determine if certain attention-allocation patterns are associated with effective cognitive-agility performance, and 2) to investigate the effects of stress on cognitive agility as measured by attention allocation. The second thesis by Lindale Fredrick (Lieutenant, USN) is titled "Cognitive Agility Measurement of Military Decision Makers" and analyzes the Make Goal data and psychological tests to determine if elements of cognitive agility measured in Make Goal can be validated by the corresponding psychological tests in the experiment.

Chapter 3 – Results

3.1. Result Summary from Thesis #1 (An Exploration of Cognitive Agility as Quantified by Attention Allocation in a Complex Environment)

This thesis yielded few significant results, but did show the value of collecting eye-tracking data during human experiments. The post-test survey revealed that participants of the stress group did report a statistically significant degree of higher stress than the control group. However, the overall stress level of both groups was not very high and it can be concluded that this military decision game does not replicate the high stress decision making tasks that are common in more demanding military jobs. Additionally, there was no significant difference in Make Goal performance between the stress group and control group. This was determined to be because the control group was not entirely stress free.

It was difficult to determine if the effects of stress could be also measured through eye-tracking data. Eye-tracking data proved to be useful when measuring Make Goal performance after dividing the subjects into high and low performing groups. Eye-tracking data showed that high performing groups spent a larger percentage of their time focused on the optimal game return on investment (ROI) metric which was the Missed Goal (regret) metric, (see Figure 6). The low performing group spent a significantly larger percentage of their time focused on the least valuable ROI metric, which was the Turns Remaining/Progress ROI. See Appendix C for the complete thesis and corresponding analysis and results.

3.2. Results Summary from Thesis #2 (Cognitive Agility Measurement of Military Decision Makers)

At this time this thesis is not complete; however, preliminary results show that subject performance on the Stroop Test, Go/No Go Paradigm and the Alternate Use Test was near the ceiling:

- Stroop test: Average of 98% correct.
- Go- No Go test: Average of 97% correct.
- Alternate Use test: Average of 17.65 items.

This is important because it will likely mean that psychological tests will not be able to be correlated to Make Goal performance or metrics, which was a key part of the experimental design. This is also important because it suggests that some traditional psychological tests are not sensitive enough to measure cognitive agility in military decision makers or at least officers attending NPS due to the homogeneity of this group.

This page left intentionally blank.

Chapter 4 – Conclusion and Recommendations

4.1. Conclusion

This project represents an effort to use interactive computer gaming to; (i) detect and measure cognitive agility in military leaders and (ii) validate it using respected psychological tests. Up to this point the results are inconclusive on whether or not we can actually measure cognitive agility using Make Goal and our chosen psychological tests. However, there are some findings that are valuable in the study of military cognitive function.

- Military decision makers are high performers and resilient to basic forms of stress. The traditional psychological tests used in this study are not sensitive enough to detect significant differences in the levels of cognitive agility of the selected military officers. Additionally, Make Goal's administrative decision task characteristic is not stressful enough to have an effect on the test subjects, despite ever changing rules, turn timers and in the case of the stress group, verbal distraction enacted by the test administrator. This is supported by the fact that there was no statistically significant difference in performance between the stress group and control group.
- Eye-tracking data is valuable when evaluating military officers' performance. Eye-tracking data from this study shows that top performers focus on optimal game ROI, while low performers do not. This could be useful in future research on military leader cognition or could be used to develop training doctrine. Often organizations know who their top and bottom performers are. With this knowledge, eye-tracking data could be collected from these top and bottom performers during critical tasks (e.g. tank gunnery, squad maneuvers, battalion command during combat operations). This data could show what information is most important for different tasks, under different circumstances and be utilized in future training doctrine. For instance, it could show what the best search patterns are, and what information is most important to a high performing tank gunner, or it could show what information is most important to a battalion commander during combat operations.

It is this study team's hope that once the second thesis is complete and analysis of the turn-by-turn Make Goal data is done that there will be some additional significant and interesting findings. It is very possible that this data will reveal efficient search patterns and officers with higher levels of rule change sensitivity in the top performance group.

4.2. Recommendations

This study team recommends that new tools be developed in order to measure cognitive agility in military decision makers. Selected psychological tests were not sensitive enough to measure the components of cognitive agility in this highly uniform subpopulation. A military computer game designed to measure cognitive agility will likely need to be more complicated and should recreate a tactical environment rather than the administrative environment featured in Make Goal. Additionally, other forms of more extreme stress should be introduced into the Stress group in order to create more variance in performance between the Stress group and Control group.

Eye-tracking data proved to be valuable and we recommend that it be used in future human domain research. We believe that its potential goes beyond cognitive agility and should extend to leader performance in tactical situations to include battalion and brigade level tactical operation center exercises.

Appendix A – **References**

1. Barron, F. (1988). Putting creativity to work. In R. J. Sternberg (Ed.), The nature of creativity (pp. 76–98). New York, NY: Cambridge University Press.
2. Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1–3):7–15. http://dx.doi.org/10.1016/0010-0277(94)90018-3.
3. Chamorro-Premuzic, T. (2006). Creativity versus conscientiousness: Which is a better predictor of student performance? *Applied Cognitive Psychology*, 20, 521–531.
4. Good, D. (2009). Explorations of Cognitive Agility: A Real Time Adaptive Capacity. Case Western Reserve University.
5. Keilp, J. G., Sackeim, H. A., & Mann, J. J. (2005). Correlates of trait impulsiveness in performance measures and neuropsychological tests. *Psychiatry Research*, 135, 191–201.
6. MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, 109, 163–203.
7. Mariscal, M. (2017). An Exploration of Cognitive Agility as Quantified by Attention Allocation in a Complex Environment. Monterey, CA: Naval Postgraduate School.
8. Martindale, C., Anderson, K., Moore, K., & West, A. N. (1996). Creativity, oversensitivity, and rate of habituation. *Personality and Individual Differences*, 20, 423–427.
9. Nesbitt, P., Kennedy, Q., Alt, J., Fricker, R. (2015). Iowa Gambling Task Modified for Military Domain. *Military Psychology,* Vol 27, No. 4, 252-260.
10. Steingroever, H., Wetzels, R., Horstmann, A., Neumann, J., & Wagenmakers, E. J. (2013). Performance of healthy participants on the Iowa Gambling Task. *Psychological Assessment, 25,* 180 –193. http://dx.doi.org/10.1037/a0029929.
11. Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 12, 643–662.
12. United States Army (USA) Combined Arms Center (CAC), (2014). The Human Dimension White Paper: A Framework for Optimizing Human Performance.
13. Zimmermann, P., & Fimm, B. (2000). Test for attentional performance (TAP). Herzogenrath: PSYTEST.

This page left intentionally blank.

Appendix B – Glossary

AUT	Alternate Use Test
CAC	Combined Arms Center
IGT	Iowa Gambling Task
IRB	Internal Review Board
LOE	Line-of-Effort
MTRY	Monterey
NPS	Naval Postgraduate School
ROI	Return on Investment
TRAC	TRADOC Analysis Center
USA	United States Army
USN	United States Navy
WCST	Wisconsin Card Sorting Test

This page left intentionally blank.

Appendix C - **Thesis #1**

AN EXPLORATION OF COGNITIVE AGILITY AS QUANTIFIED BY ATTENTION ALLOCATION IN A COMPLEX ENVIRONMENT

By: Miriam C. Mariscal

The attached thesis supports the Army's mission to establish a measure for cognitive agility in soldiers. It examines attention-allocation patterns as quantified by eye-tracking data collected while subjects played a military-relevant cognitive agility computer game (Make Goal), to determine whether certain patterns are associated with effective performance. It also investigates the effects of stress on cognitive agility as measured by attention allocation.

Methods: Forty military officers were randomly assigned to a stress or control group. Stress level was manipulated by timed turns and experimenter behavior.

Results: Eye-tracking data was analyzed in terms of regions of interest on which subjects focused their cognitive workload. Results were analyzed by stress and control group and top and bottom ten performers. The stress and control groups showed similar attention-allocation patterns. The high performers attended more to the important information and made more optimal selections than poor performers.

Discussion: Results are discussed in the context of the Yerkes-Dodson stress model. Eye-tracking data revealed attention-allocation patterns associated with higher performance. In order to better detect the impact of stress on Cognitive Agility, an experiment that includes a wider range of stress levels is needed.

This page left intentionally blank.

NAVAL
POSTGRADUATE
SCHOOL

MONTEREY, CALIFORNIA

THESIS

AN EXPLORATION OF COGNITIVE AGILITY AS QUANTIFIED BY ATTENTION ALLOCATION IN A COMPLEX ENVIRONMENT

by

Miriam C. Mariscal

March 2017

Thesis Advisor:	Quinn Kennedy
Second Reader:	Samuel Buttrey

Approved for public release. Distribution is unlimited.

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE	*Form Approved OMB No. 0704-0188*

1. AGENCY USE ONLY ***(Leave blank)***	**2. REPORT DATE** March 2017	**3. REPORT TYPE AND DATES COVERED** Master's thesis

4. TITLE AND SUBTITLE AN EXPLORATION OF COGNITIVE AGILITY AS QUANTIFIED BY ATTENTION ALLOCATION IN A COMPLEX ENVIRONMENT	**5. FUNDING NUMBERS**
6. AUTHOR(S) Miriam C. Mariscal	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000	**8. PERFORMING ORGANIZATION REPORT NUMBER**
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A	**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB number 2016.0048-IR-EP7.

12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release. Distribution is unlimited.	**12b. DISTRIBUTION CODE**

13. Abstract (maximum 200 words)

This thesis supports the Army's mission to establish a measure for cognitive agility in soldiers. It examines attention-allocation patterns as quantified by eye-tracking data collected while subjects played a military-relevant cognitive agility computer game (Make Goal), to determine whether certain patterns are associated with effective performance. It also investigates the effects of stress on cognitive agility as measured by attention allocation.

Methods: Forty military officers were randomly assigned to a stress or control group. Stress level was manipulated by timed turns and experimenter behavior.

Results: Eye-tracking data was analyzed in terms of regions of interest on which subjects focused their cognitive workload. Results were analyzed by stress and control group and top and bottom ten performers. The stress and control groups showed similar attention-allocation patterns. The high performers attended more to the important information and made more optimal selections than poor performers.

Discussion: Results are discussed in the context of the Yerkes-Dodson stress model. Eye-tracking data revealed attention-allocation patterns associated with higher performance. In order to better detect the impact of stress on Cognitive Agility, an experiment that includes a wider range of stress levels is needed.

14. SUBJECT TERMS eye tracking, cognitive agility, attention allocation patterns, stress	**15. NUMBER OF PAGES** 81
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	**18. SECURITY CLASSIFICATION OF THIS PAGE**	**19. SECURITY CLASSIFICATION OF ABSTRACT**	**20. LIMITATION OF ABSTRACT**
Unclassified	Unclassified	Unclassified	UU

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

THIS PAGE INTENTIONALLY LEFT BLANK

AN EXPLORATION OF COGNITIVE AGILITY AS QUANTIFIED BY ATTENTION ALLOCATION IN A COMPLEX ENVIRONMENT

Miriam C. Mariscal
Lieutenant, United States Navy
B.S., United States Naval Academy, 2012

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
March 2017

Approved by: Quinn Kennedy, Ph.D.
Thesis Advisor

Samuel Buttrey, Ph.D.
Second Reader

Patricia A. Jacobs, Ph.D.
Chair, Department of Operations Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

This thesis supports the Army's mission to establish a measure for cognitive agility in soldiers. It examines attention-allocation patterns as quantified by eye-tracking data collected while subjects played a military-relevant cognitive agility computer game (Make Goal), to determine whether certain patterns are associated with effective performance. It also investigates the effects of stress on cognitive agility as measured by attention allocation.

Methods: Forty military officers were randomly assigned to a stress or control group. Stress level was manipulated by timed turns and experimenter behavior.

Results: Eye-tracking data was analyzed in terms of regions of interest on which subjects focused their cognitive workload. Results were analyzed by stress and control group and top and bottom ten performers. The stress and control groups showed similar attention-allocation patterns. The high performers attended more to the important information and made more optimal selections than poor performers.

Discussion: Results are discussed in the context of the Yerkes-Dodson stress model. Eye-tracking data revealed attention-allocation patterns associated with higher performance. In order to better detect the impact of stress on Cognitive Agility, an experiment that includes a wider range of stress levels is needed.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

AUT	Alternate Use Test
IRB	Institutional Review Board
ms	Millisecond
NRC	Navy Recruiting Command
QMA	Qualified for Military Application
ROI	Region of Interest
TRAC	TRADOC Analysis Center
TRADOC	United States Army Training and Doctrine Command

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

The Army has identified cognitive dominance as being essential to maintaining superiority over adversaries (United States Army Combined Arms Center, 2014). However, the Army lacks a standardized means of identifying and quantifying cognitive agility, and this thesis is part of a larger effort by TRADOC to develop just such a measure. In support of this effort, this thesis has a two-fold goal. First, it examines attention-allocation patterns demonstrated by subjects during an experiment in order to determine if certain attention-allocation patterns are associated with effective cognitive-agility performance. Second, this thesis aims to investigate the effects of stress on cognitive agility as measured by attention allocation. Eye-tracking data was used to measure attention allocation, as it provides a non-intrusive, objective and quantifiable means of gathering information on subjects' cognitive processes, as well as the effects of stress on decision making (Sullivan, Yang, Day, & Kennedy, 2011; Vine, Uiga, Lavric, Moore, Tsaneva-Atanasova, & Wilson, 2014). This thesis is an exploratory study, and it is hoped that any information gleaned will be of use in directing further study in this area.

This thesis examines attention-allocation patterns demonstrated by subjects while they played a military-relevant cognitive agility computer game called Make Goal. Make Goal requires subjects to assess and use given items of information about a recruiting region (i.e., the number of individuals qualified for military service in that region). Subjects must then decide which items of information are important and use them to make optimal decisions regarding recruiter placement to gain the maximum number of recruits within one hundred timed trials. The relative importance of these information items changes throughout the game, requiring subjects to detect the change and shift strategy accordingly. Attention allocation was measured by collecting eye-tracking data as each subject played the game. The game screen was split into various regions of interest (ROIs) so that the eye-tracking data could be analyzed to determine at which information a subject was directing his or her attention during the game and for how long. Cognitive workload data also was collected via measurement of pupil dilation.

The subjects in this experiment consisted of forty military officer volunteers who were randomly assigned to either the stress or control group. Those in the stress group were subjected to a social stress induction, in which the experimenter monitored them closely while they played the game, with periodic comments regarding their game performance. In contrast, those in the control group received no such monitoring by the experimenter. It was expected that the stress group would either do much better or much worse than the control group, according to the Yerkes-Dodson inverted U relationship between stress and performance. The Yerkes-Dodson model postulates that both very high and very low levels of stress can negatively impact performance, while moderate stress can actually improve performance (Yerkes & Dodson, 1908). In order to determine if certain attention-allocation patterns are associated with effective cognitive agility, performance eye-tracking data also was examined among the best and worst performers.

Analysis of the results (both the Make Goal Game results and the eye-tracking measures) showed almost no significant differences in attention-allocation patterns and performance between the stress and control groups. Only one significant difference was found; the control group looked more often at one particular ROI, Turn Remaining/ Progress (which was not strategically important to performance). However, several differences were found between the high- and low-performing groups. As might be expected, high performers were better able to identify the most relevant pieces of information and use this information to make optimal recruiting decisions.

The precise role of stress in affecting performance in this study was not clear from the analysis, as almost no significant differences were found in attention-allocation patterns between the stress and control groups. Subjective levels of stress were reported in a post-task survey. While the stress group reported the experimenter's behavior as being more stressful than the control group, the control group was not entirely free from stress; eleven out of the twenty control-group subjects reported the game timer as being stressful. The top ten high performers and bottom ten low performers both shared mixes of subjects from both the stress and control groups. Interestingly, though, the high performers reported the timer as being significantly more stressful than the low performers did. This finding is consistent with the Yerkes-Dodson relationship between stress and performance, in which moderate stress is associated with improved performance. In order to better analyze the effects of stress, some experimental design improvements could be made, including removal of the timer from the game interface for the control group so that the control group encounters no stressors. Perhaps an EEG or skin conductance test would be useful, in order to provide a more specific classification of stress levels than that provided by the eye tracking and subjects' self-reports. A more precise and standardized classification of stress levels would provide a better test of the Yerkes-Dodson model.

In conclusion, the ultimate goal of this TRADOC study is first, to identify cognitively agile military personnel and second, to investigate the effects of stress on cognitive agility. This thesis found that eye tracking did, in fact, provide insight into the decision-making processes of more cognitively agile subjects; it revealed that they found and then focused on more relevant information and therefore performed better in the game. Although the results comparing the stress and control groups were not what was expected, the fact that eye-tracking data proved useful in identifying attention-allocation patterns between high and low performers shows promise for future work.

References

Sullivan, J., Yang, J. H., Day, M., & Kennedy, Q. (2011). Training simulation for helicopter navigation by characterizing visual scan patterns. *Aviation Space and Environmental Medicine 82*(9), 1–8.

United States Army Combined Arms Center. (2014). *The human dimension white paper*: *a framework for optimizing human performance.*

Vine, S. J., Uiga, L., Lavric, A., Moore, L. J., Tsaneva-Atanasova, K., & Wilson, M. R. (2014). Individual reactions to stress predict performance during a critical aviation incident. *Anxiety, Stress and Coping 4*(28), 467–477.

Yerkes, R.M., & Dodson, J.D. (1908). The relation of strength of stimulus to rapidity of habit information. *Journal of Comparative Neurology of Psychology, 18*, 459–482.

THIS PAGE INTENTIONALLY LEFT BLANK

ACKNOWLEDGMENTS

I would like to thank everyone who helped and supported this project and my thesis. First and foremost is my advisor, Dr. Kennedy, for all of her help and guidance in this thesis from start to finish, and for being one of the most responsive and dedicated people with whom I have worked.

I would like to thank my second reader, Dr. Buttrey, for all of his help getting through the daunting data analysis involved in this project.

I would also like to thank Major Adam Haupt, USA; Mr. Jimmy Liberato; and the techs from the MOVES department who lent support during the experiments, as well as the forty volunteers who participated.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. BACKGROUND

In the Army's Force 2025 and Beyond planning process, cognitive dominance was identified as an essential component to maintaining a "decisive edge" over adversaries (United States Army Combined Arms Center, 2014). In support of this emphasis on cognitive agility, the TRADOC Analysis Center has undertaken a study seeking to (1) operationalize and measure cognitive agility in a military context and (2) determine if psychological measures of cognitive flexibility, focused attention and creativity correlate with outcome measures of a military cognitive-agility computer game. Although aspects of cognitive agility have been studied in the civilian sector (Good, 2014), most of these studies lack the stress and high stakes of a military environment. Therefore, the TRADOC study, which is the focus of this thesis, aims to create a military-relevant measure of cognitive agility that replicates some of the stressors typically experienced by military decision makers. The ultimate goal of this body of work is first, to identify cognitively agile military personnel and second, to investigate the effects of stress on cognitive agility. Thus, it is important to understand the attention-allocation patterns used by the cognitively agile, because these methods can uncover underlying cognitive strategies (Sullivan, Yang, Day, & Kennedy, 2011). This thesis supports the Army's goal by examining attention-allocation patterns as they are quantified by eye-tracking data, seeking to determine if certain attention-allocation patterns are associated with effective cognitive-agility performance, and by investigating the effects of stress on cognitive agility as indicated by attention allocation.

The means used in this study to gather information about attention allocation and underlying strategy was eye tracking. Eye-tracking data is objective and quantifiable and can provide insight into such cognitive processes as decision making (Spivey, Richardson & Fitneva, 2004). This thesis focuses on a military relevant cognitive-agility computer game called "Make Goal," in which subjects have to allocate recruiters to regions in order to meet recruiting goals within a certain number of trials (Alt, Appleget, Buttrey, Kennedy, Sciarini, & Johnson, 2015). In order to make optimal recruiting decisions, subjects must rely on particular pieces of information such as the map of the recruiting districts, amount of time remaining, overall goal, number of personnel, and output. Periodically, the relative importance of each piece of information changes, resulting in a need to shift one's strategy. Thus, the game measures how well subjects attend to feedback and how quickly they realize that they need to change strategy.

Subjects played Make Goal under one of two conditions. A stress group had the experimenter closely monitoring them while they played the game (experimenter-induced stress), and the control group was not subject to this stress. Eye-tracking data was collected on each subject during the game. During the original study of the Make Goal game, the eye tracking results indicated the ability to discern subjects' transition from exploration to exploitation of the game (Alt et al., 2015). It is hoped that similar insights can be gleaned from this study by comparing gaming performance of the subjects and their eye-tracking data, to establish whether patterns emerge indicating that more successful subjects have different eye-tracking patterns from the less successful.

The impact of stress on cognitive performance is well documented (Eysenck, Derakshan, Santos, & Calvo, 2007), and the measurability of this impact by using eye-tracking data has been established. According to Wilson (2012) as cited in Vine, Uiga, Lavric, Moore, Tsaneva-Atanasova and Wilson (2014), anxiety is known to cause disruption in a focused gaze leading to "more, shorter fixations to a variety of locations, and (subjects) are unable to maintain the long, target-focused fixations important for the planning and control of movement" (p. 469). Thus, previous work suggests that stress should influence attention allocation during the Make Goal Game.

B. LITERATURE REVIEW

1. Eye Tracking

As mentioned above, the use of eye tracking in order to quantify cognitive processes has been established (Poole, Ball & Phillips, 2005; Spivey et al., 2004; Sullivan et al. 2011). However, the majority of the relevant studies using eye-tracking metrics as a means of quantifying cognitive processes have involved pilots (Bellenkes, Wickens & Kramer, 1997; Kasarskis, Stehwien, Hickox, Aretz, & Wickens, 2001; Sarter, Mumaw &

Wickens, 2007; Sullivan et al., 2011; Vine et al., 2014) or drivers (Konstantopoulos, Chapman & Crundall, 2010; Maltz & Shinar, 1999; McPhee, Scialfa, Dennis, Ho, & Caird, 2004; Wikman & Summala, 2005). These studies involve subjects performing a task with which they are familiar and subjects who can be categorized on a spectrum from novice to experienced. Therefore, the eye-tracking data has been used to compare the differences in the decision-making processes of subjects across this spectrum of varied experience.

The use of eye tracking in this thesis differs fundamentally from these other studies. Rather than trying to understand how subjects' cognitive processes vary with or relate to their experience level, it seeks to understand, or at least quantify, the cognitive process itself, according to how subjects performed on a decision-making task with which they had no prior experience. Studies with pilots have found that better decisions are made by the more experienced pilots due to the fact that they are more efficient and effective in how they allocate their attention (Schriver, Morrow, Wickens & Talleur, 2008; Sullivan et al., 2011).

This thesis investigated whether attention-allocation patterns can identify the more cognitively agile. It examined the eye-tracking data of subjects as they played Make Goal, which required them to shift strategies throughout the game's duration. The thesis examined whether attention-allocation patterns differed between high and low performers and between the stress or control group. With regard to the eye tracking, this study focuses on fixation-duration data as well as workload (indicated by pupil dilation). These measures are typically used in correlating eye tracking and cognitive processes. The three eye-tracking measures used most frequently in cognitive analysis are saccades, fixations and pupil dilation. Although this study does not focus on saccades, per se, it is worth noting what a saccade is because fixations, which this study does examine, are defined by saccades. A saccade is defined as the quick movement of the gaze lasting between 30 and 50 milliseconds going from consecutive points of interest, while a fixation is defined as the time between saccades, usually lasting 200–300 milliseconds (McCarley & Kramer 2007). Fixations, which are one of the primary areas of interest in this study, are particularly important as it is during fixations that a subject acquires information

(McCarley & Kramer 2007). "A number of studies have shown that gaze duration is sensitive to the information value of the item being fixated, with items that are unexpected or semantically inconsistent with their context receiving longer gazes than items that are expected within the context (e.g., Friedman, 1979; Loftus & Mackworth, 1978)," as cited by McCarley and Kramer (2007, p. 97). Furthermore, "a number of studies have shown that gaze duration (what we in this study are calling a fixation) is sensitive to the information value of the item being fixated" (McCarley & Kramer 2007, p. 97). The link between pupil dilation and cognitive processes is established; generally, a subject's pupils will dilate in accord with a higher cognitive workload (Kahneman, 1973; Marshall, 2007).

The study focuses on subjects' visual fixations on regions of interest (ROI) within the Make Goal Task. In the context of the study, an ROI is simply an area on the computer screen that provides meaningful information to the subject. The examination of the breakdown of subjects' fixations on different ROIs on the screen can yield information about what they considered important, or what given information they were using to inform their decisions. Finally, the link between pupil dilation and cognitive processes is established (Marshall, 2007). Specifically, a subject's pupils will dilate when they experience an increase in cognitive workload. (Kahneman, 1973). Thus, larger dilation is positively associated with higher cognitive workload.

2. Stress

Stress is an important aspect of this study because it is an intrinsic part of the military environment; understanding its effects on the cognitive agility of decision makers is one of the Army's key objectives in this study. There are varying definitions of stress; the one used for this thesis is as defined by Stokes and Kite (2001). Stress is "the result of a mismatch between individuals' perceptions of the demands of the task or situation and their perceptions of the resources for coping with them" (Stokes & Kite, 2001, p. 116). Just as there are varying definitions of stress itself, the precise impact of stress on cognitive performance is also the subject of study and debate. It seems to be the finding of the majority of studies that stress adversely impacts cognitive performance, from memory to decision making to attention and perceptual-motor performance (Ashcraft & Kirk, 2001; Ashcraft, 2002; Braunstein-Bercovitz, Dimentman-Ashkenazi & Lubow, 2001; Enander 1989; Eysenck et al., 2007; Friedland, Keinan & Regev, 1992; Lieberman, Bathalon, Falco, Georgelis, Morgan, Niro & Tharion, 2002; Matthews & Desmond, 2002; Ozel, 2001; Robert & Hockey 1997; Van Galen & van Huygevoort, 2000).

There are other studies that seem to indicate that some types of stress may actually improve cognitive performance. For example, a study conducted by Lavine, Sibert, Gokturk, and Dickens (2002), wherein subjects were exposed to stress in the form of random noise bursts, found their performance on a vigilance task improved. Similar findings were obtained in a study by Kirk and Hecht (1963), also involving noise stress. Another study conducted by Ozel (2001) involved fire-fighters' decision-making processes in egress from a fire. Ozel (2001) found that when exposed to stress in the form of time and fire, low amounts of stress resulted in better decisions regarding when to exit, whereas higher stress levels resulted in degraded decision making. As cited by Staal (2004), Chajut and Algom (2003) found that in some cases "stress reduced interference and improved selective attention" (Staal, 2004, p. 42). Braunstein-Bercovitz's study (2003), cited by Staal (2004) "demonstrates that there is an interaction effect between workload and psychological stress. When workload is relatively low and stress is high, the selective attention effect is present (negative priming is attenuated). On the other hand, when both workload and stress are high, support for the selective attention hypothesis diminishes and the negative priming effect is strong" (Staal, 2004, p. 41). Negative priming refers to the phenomenon wherein after making a conscious effort to ignore a stimulus, subsequent attempts to process that stimulus will be impaired (Tipper, 1985). For example, if a subject is told to focus on picking a green pen out of a bunch of blue pens, and afterwards is told to focus on picking out a blue pen, the subject's initial difficulty in switching from focusing on the green pen to the blue pen would be an example of negative priming.

The above findings suggest that stress can be beneficial or detrimental to cognitive processes, depending on the type and level of stress, the task being measured and also the individual subject, e.g., subjects already nervous or anxious by nature tend to show degraded cognitive performance under stress (Wofford, Goodwin & Daly, 1999; Wofford, 2001; Wofford & Goodwin, 2002). This non-monotonic relationship between stress and performance was first proposed by Yerkes and Dodson (1908). This model proposes that the relationship between stress and performance can be modeled by an inverted U, wherein very low and very high levels of stress are detrimental to performance, but a moderate level of stress corresponds to better performance. Quantifying the effects of stress on an individual subject still presents a challenge because the same stimuli can be perceived as stressful by some individuals and not stressful by others, and even if a stimulus is perceived as stressful, individual reactions can still differ (Cerin, Szabo, Hunt, & Williams, 2000). Thus, eye tracking provides one method for quantifying individual differences in stress response across subjects. The measurability of this impact by using eye-tracking data has likewise been established. According to Wilson (2012) as cited in Vine et al. (2014), anxiety is known to cause disruption in a focused gaze leading to "more, shorter fixations to a variety of locations, and (subjects) are unable to maintain the long, target-focused fixations important for the planning and control of movement" (Vine et al., 2014, p. 469). Studies conducted on surgeons, athletes, and pilots have demonstrated that disruptions in visual fixations have been linked to degraded performance (Vine et al., 2014). For example, among pilots who completed either a stressful or non-stressful scenario, the pilots in the stressful scenario showed more disruptions in visual fixations, "higher search rates and increased randomness in scanning behavior" (Vine et al., 2014, p. 12). Thus, previous work suggests that stress should influence attention allocation during the Make Goal Game. Specifically, we would expect to see that subjects in the stress condition will exhibit more, shorter fixations and greater workload as indicated by pupil size.

3. Literature Review Summary

The current literature provides a background and starting point for this study. It indicates that stress can be a subjective influence, depending upon the individual, and that it is likely to have some effect, either positive or negative, upon the subject's cognitive processes. The social induction stress used in this thesis would most likely correspond to a moderate level of stress. Based on the literature, and particularly, the Yerkes-Dodson model of stress and performance, improved performance might be expected from the stress group. Furthermore, eye tracking should be a reliable means of quantifying the effects of stress upon subjects' cognitive processes. We expect to see differences in the pupil dilation and fixations exhibited by the two groups in this study.

C. OVERVIEW OF THESIS EXPERIMENT AND RESEARCH QUESTIONS

As previously mentioned, this thesis explores part of a study conducted by TRAC Monterey in order to address U.S. Army objectives regarding identifying and quantifying cognitive agility in its personnel. The study involved several traditional psychological tests of cognitive agility, including the Stroop Test (Stroop, 1935), the Go No Go Test (Donders, 1969), and the Alternate Use Test (Guilford, Christensen, Merrifield & Wilson, 1978), with the addition of the Make Goal recruiting scenario game, during which subjects were part of either a control or stress group, and eye-tracking data was collected. Each of these components will be addressed further in the methods chapter, but this thesis focuses specifically on the Make Goal task and analysis of the accompanying eye-tracking data.

This is a preliminary study conducting exploratory analysis, and as such does not have any formal hypotheses; however, the analysis has the following objectives:

1. To determine whether there is an association between attention-allocation patterns and performance on the Make Goal computer game.

2. To determine whether social stress affects attention allocation patterns.

THIS PAGE INTENTIONALLY LEFT BLANK

II. METHODOLOGY

A. PARTICIPANTS

The subjects in this study consisted of forty military officers, both U.S. and international, from all service branches. See Table 1 for demographic characteristics of the sample. Subjects were all volunteers recruited from among the students and staff at the Naval Postgraduate School. Recruitment was conducted by word of mouth, mass emails, and the posting of an advertisement on the school's daily muster page, where students are required to check a list of daily announcements. As subjects responded to participate, they were randomly assigned to one of two groups: a stress group or a control group. Statistical analysis demonstrated that there are no statistically significant differences between the stress and control groups in terms of service, rank, years in service, age or gender.

Table 1. Subject Demographic Data

Demographic Data	Control Group	Stress Group
Service	US Army: 4 US Marine Corps: 6 US Navy: 10	US Army: 4 US Marine Corps: 3 US Navy: 12 International: 1
Rank	O5: 1 O4: 9 O3: 9 O1: 1	O4: 7 O3: 9 O2:4
Years in Service Mean (sd)	11.05 (6.14)	9 (4.76)
Age Mean (sd)	33.7 (6.28)	30.7 (4.49)
Gender	14 male 6 female	10 male 10 female

B. MEASURES

1. Make Goal Task

This task is a recruiting scenario game in which subjects are given one hundred timed turns in which to place recruiters in various districts to generate enough recruits in order to meet a given goal of seven hundred twenty-five recruits. The time in which to make a move on each turn becomes progressively shorter over the course of the game: subjects begin with one hundred seconds per turn and end with only fifteen seconds per turn. Certain demographic data (labeled in Figure 1 as regional data categories) are given in a column on the left-hand part of the screen; subjects can use this demographic data to try and ascertain the best possible district in which to place their recruiter on each turn. The demographic data include the number of youths between the ages of seventeen and twenty-four, the number of individuals qualified for military service and the number of members of minority groups. Subjects can click on any of these attributes to see the districts on the map highlighted according to these attributes (i.e., the region with the most youth between the ages of seventeen and twenty-four would be dark green, the region with second most would be light green, etc.) On the right side of the screen the names of each of the recruiting districts are listed (labeled in Figure 1 as "station information based on regional data"). During the game the regions are not labeled on the map itself; subjects can click on station information to view each district highlighted on the map. Subjects must ascertain how each of these attributes contributes to the number of recruits generated in order to make optimal recruiter placement decisions. After submitting each turn by clicking on the next turn button, subjects receive immediate feedback regarding the number of recruits generated on the last turn as well as by how much they missed their goal, if at all. The relative importance of each piece of information (which influences the "optimal selection") changes six times over the course of the game, resulting in a need for the subject to detect the change and accordingly shift strategy:

- Turns 1–20: Focusing on the data attribute diversity (number of members of minority groups) will lead to the optimal number of recruits.

- Turns 21–38: The most salient data attribute for these turns is population qualified for military service.

- Turns 39–55: The most salient data attribute for these turns is diversity (number of members of minority groups).

- Turns 56–71: The most salient data attribute for these turns is the number of youth between 17 and 24.

- Turns 72–87: The most salient data attribute for these turns is the population qualified for military service.

- Turns 88–100: The most salient data attribute for these turns is diversity (number of members of minority groups).

(Appendix A shows a table listing the optimal station selections and how these changed throughout the game). As shown in Figure 2, subjects also receive information regarding the amount of time left in that turn, the total number of recruits (accessions) they have recruited so far, and the number of turns remaining. For the purposes of this thesis, the main outcome measure was total recruits. Figure 1 depicts the Make Goal game screen and Table 2 lists the relevant Make Goal variables.

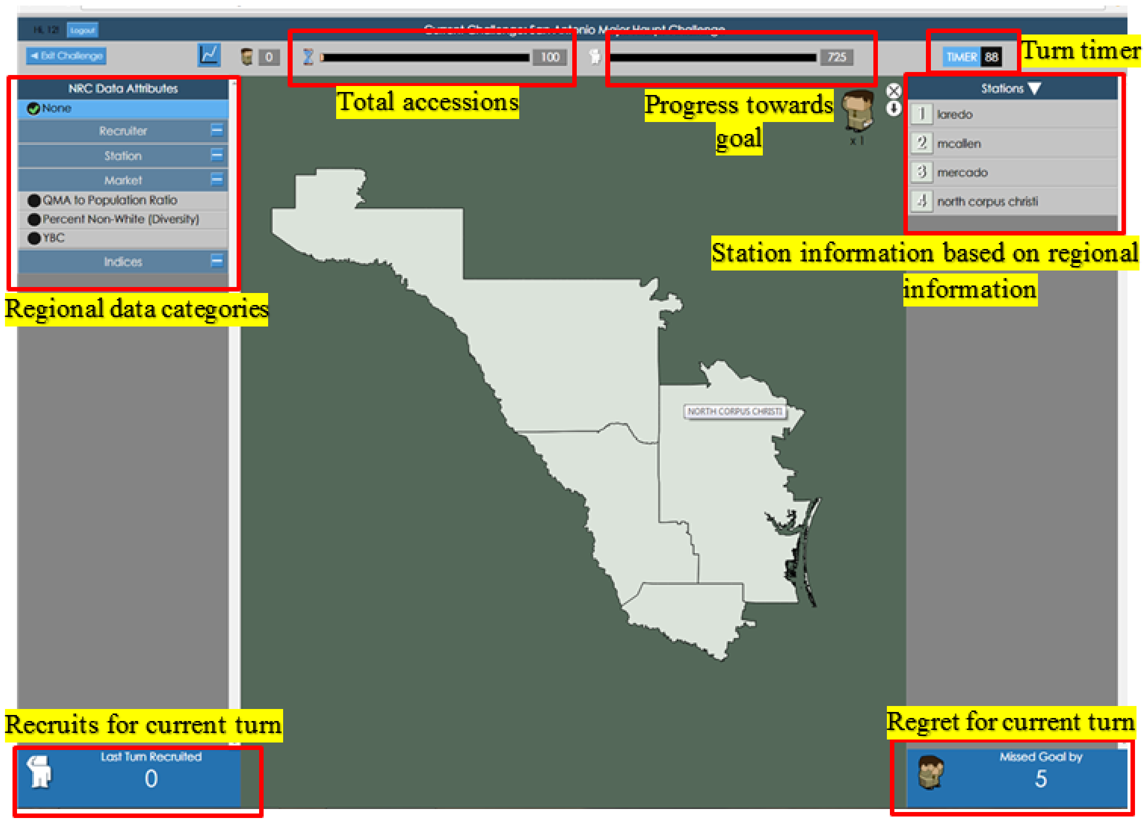


Figure 1. Make Goal Screen

Table 2. Make Goal Game Terminology Definitions

Make Goal Measure	Definition
Total Accessions	Total number of recruits gained during the Make Goal game
Regret	Total number of possible recruits missed during the Make Goal game
% Optimal Selection	Once a subject makes an optimal selection (see Chapter III, section B, part 1 for optimal selection definition) the percentage of trials for which the subject continues to stay with that selection while it remains optimal
Time	Duration of time taken to complete the Make Goal game

2. Stress Manipulation

There were two sources of stress manipulated in this experiment: experimenter-induced stress and timer stress.

a. Experimenter-Induced Stress

The randomly assigned members of the stress group played Make Goal with the experimenter observing them closely and making comments as they played the game. A script was followed by the experimenter to ensure consistency for all subjects. At the following intervals the following comments were made:

- Move 35: “Are you sure you want to make that move?”
- Move 50: “You might want to take a look at the timer.”
- Move 70: “Usually people are doing a little better by this point.”
- Move 85: “Might want to check on that timer again.”
- Move 90: “It doesn’t look like you’re doing too well.”

The randomly assigned members of the control group played Make Goal without the experimenter observing them closely or making any comments on their performance.

b. Timer Stress

The Make Goal screen has a timer in the upper right-hand corner of the screen (see Figure 1). The time in which to make a move on each turn becomes progressively shorter over the course of the game: subjects begin with ninety seconds per turn and end with only fifteen seconds per turn. Due to constraints in game design, all subjects, both in the control and stress groups, experienced timer stress.

3. Typical Psychological Measures of Cognitive Agility

a. Stroop Test (Stroop, 1935)

In the Stroop Test, subjects are shown a series of names of colors written in different color fonts. They have to react as quickly as possible to the color in which the word is written while ignoring what the word itself reads, by hitting color coded arrow keys on the keyboard. For example, in Figure 2, the correct responses would be red and green, respectively. The test is scored by calculating the percentage of correct responses.

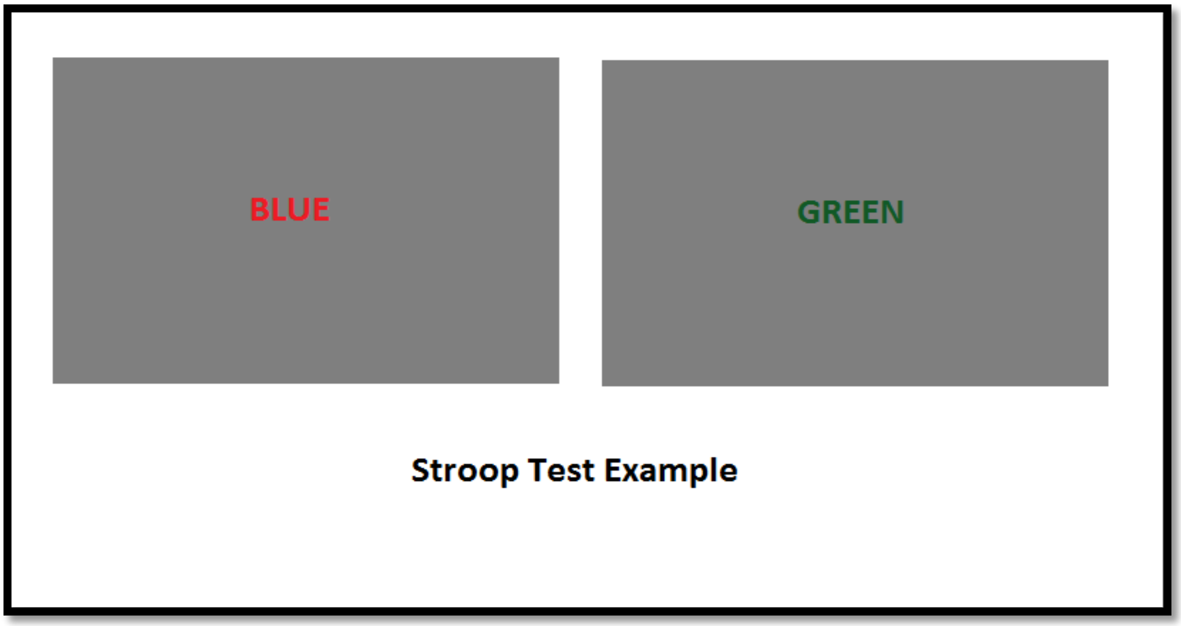


Figure 2. Stroop Test Screenshot

b. Go No Go Test (Donders, 1969)

This test is an eight-minute-long test of focused attention and reaction time, wherein subjects are shown a series of shapes that were either green or blue. If the shape is green, they have to hit the space bar as quickly as possible; otherwise, they are not supposed to do anything (see Figure 3). Thus, over time, the subject must refrain from the automatic response of hitting the space bar when they are presented with a blue shape. The test is scored by calculating the percentage of correct responses.

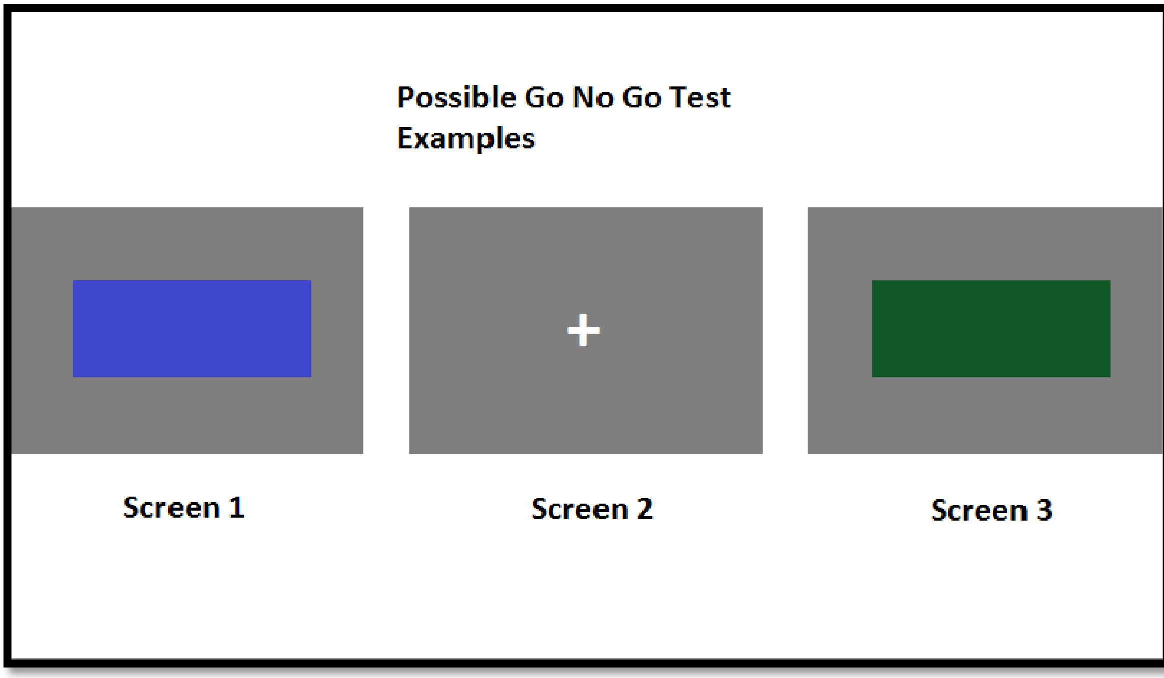


Figure 3. Go No Go Test Screenshot

c. Alternate Use Test (AUT) (Guilford, Christensen, Merrifield & Wilson, 1978)

The AUT is a creativity test in which subjects are given three minutes to list as many unique uses as they can imagine for a brick. In this study, subjects typed in their responses on the computer. The score is the total number of non-repeated responses listed.

4. Surveys

a. Demographic Survey

This survey collected information on subjects' service branch, age, time in service, recruiting experience (if any, and how long) and time spent playing computer games each week (if at all, which type).

b. Post Make Goal Accessions Survey

Subjects were asked to list everything they believed influenced the number of accessions they achieved each turn.

c. Post Task Survey

In this survey, subjects answered questions regarding their strategy during the Make Goal Task, specified which pieces of information they found most useful, and rated how stressful they found the game and various elements of the game on a 1–7 Likert scale, including the timer and experimenter's behavior. See Appendix B for a copy of the survey.

5. Eye Tracking

Table 3 lists the eye-tracking terminology used in this thesis along with brief definitions. These measures are used to quantify subjects' attention allocation while they played Make Goal.

Table 3. Eye-Tracking Terminology Definitions

Eye-Tracking Measure	Definition
Fixation	Maintaining visual gaze on a single location, typically lasting 200–300 ms.
Number of Fixations	How many times each subject fixated during the experiment.
Fixation Duration Overall	Overall average length (in seconds) of each subject's fixations during the experiment.
Fixation Duration by ROI	Overall average length (in seconds) of each subject's fixations in each ROI.
% Fixation Time by ROI	Time a subject spent fixated in each ROI as a percentage of total fixation time throughout the game duration.
Number of Fixation Shifts	How many times a subject's fixation shifted from one ROI to another during the experiment.
Fixation Shift Rate	How many times a subject's fixation shifted from one ROI to another during the experiment divided by the total number of fixations.
Workload	Measure of subject's effort according to pupil dilation, on a scale from .0001 to 1.0000 in which higher values indicate higher cognitive workload.

During the Make Goal Task, subjects' eye movements were measured and recorded using the FaceLAB eye-tracking equipment and Eye Works software. Eye-tracking data is collected 60 times per second; pupil dilation data is collected once per second. Eye movements are measured in gaze duration in each region of interest (ROI), percent of time subjects looked at each ROI, number of gaze shifts between each ROI, and workload as measured by pupil dilation throughout the task. There were six ROIs: the timer and the list of recruiting stations, the NRC data, the regret number (number of recruits by which goal was missed), the map, the number recruited on the last turn, the number of turns remaining as well as the progress being made towards making goal. (See Figure 4 for the ROI's as well as sample eye tracking data overlay.)

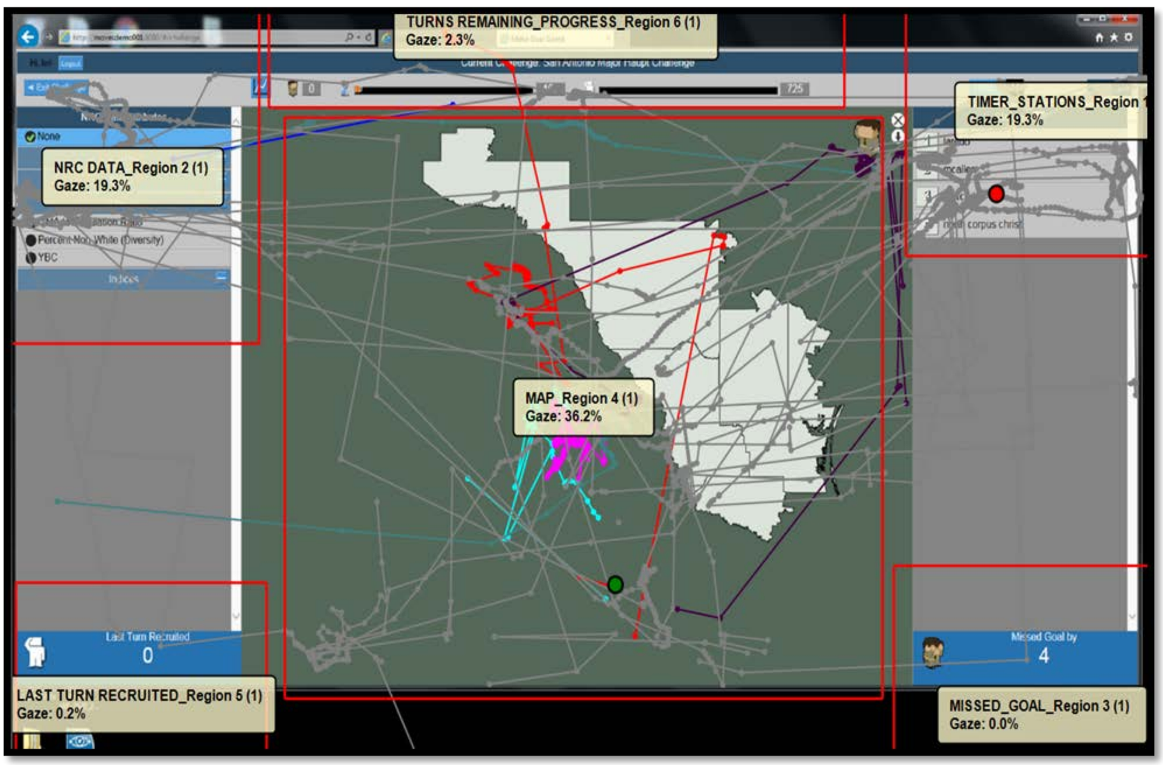


Figure 4. Make Goal Regions of Interest

C. EQUIPMENT

The equipment used in this study consists of two desktop computers with Windows 7 operating systems and two eye-tracking stereo cameras. The cameras detect reflected infrared light from a subject's face and eyes with 12 mm lenses. The cameras are calibrated to each subject to find reference points on the face (outside corners of the eyes and mouth), which enable the camera to monitor the subject's head position, direction and duration of gaze, as well as workload from pupil dilation. This information is recorded by EyeWorks record on desktop computer number two (which also runs all of the other tasks in the study). Desktop number one is used only to run FaceLAB 5.0.7 30.

D. PROCEDURES

This study was approved by the NPS IRB. The experiment typically took less than one hour. Upon arrival at the lab, the general purpose of the study was explained to subjects; the experimenter used a script to explain the tasks and task requirements to each subject. Subjects were not told whether they were in a stress or control group or what each task measured until the debrief, and they were not allowed to ask questions of the experimenter once a task was begun.

First, subjects completed the IRB-approved consent form, followed by the demographic survey, and then completed the tasks in the following order: Stroop, Go No Go and AUT. These tasks were completed for the larger TRADOC study and are not the focus of this thesis. The eye-tracking equipment was then calibrated to the subject. Once calibration was successfully accomplished, the subject completed the Make Goal Task and had his/her eye tracking recorded. Next, subjects completed the Post Make Goal Accessions Survey and the Post Task Survey. Finally, if the subject was a member of the stress group s/he was debriefed, informing him/her that s/he was a member of a stress group and that the stress induction experienced was part of the experiment. The experimenter also answered any questions the subject may have had about the study, procedures, goals and potential implications and uses of the results in developing future training Army's 2025 Vision Plan.

III. RESULTS

A. DATA PREPARATION AND STATISTICAL ANALYSIS

1. Data Preparation

The raw eye-tracking data was in the form of two spreadsheets. One spreadsheet contained the start and end time of each fixation, as well as the fixation duration by each ROI. There also is a default ROI called “empty,” which corresponds either to when the subject blinked, lost eye tracking due to looking away from the screen, or looked at non-ROI areas of the screen. The “empty” ROI contributes nothing of importance to this study so it was excluded from the analysis. The second spreadsheet consisted of the workload data. The eye-tracking software also produces a video of each subject’s eye gaze while they played Make Goal. The Make Goal Game output was in the form of a comma separated value (csv) file with all of the game data, including the number of accessions (the actual score) and regret (number of recruits missed). Regret was calculated as the difference between the total number of recruits that could possibly have been gained by selecting the optimal location, and the number actually gained in the location selected.

Due to the nature of the eye-tracking equipment, the eye tracker would begin before the subject began the Make Goal game and end after the subject had completed Make Goal. In order to remedy this timing issue and ensure that all start and end times were accurate, the video output depicting the gaze trace of each subject as they played the Make Goal game was examined for the exact time stamps corresponding to the beginning and end of the game. This data was compiled in a spreadsheet that was read into R (R Core Team 2015) along with the eye-tracking data. R was then used to obtain the averages of all of the eye tracking and Make Goal measures (number of fixations, fixation duration overall, fixation duration by each ROI, workload, number of fixation shifts, shift rate, accessions, regret, % optimal selection, and time), according to the definitions listed in Chapter II. They were all output to a master file, which was used for the statistical analysis. All of the measures in the control group had (n = 20) observations, while the stress group had (n = 19) observations, as one of the stress group subjects did not yield usable eye-tracking data.

2. Statistical Analysis Methods

Two sets of analyses between groups were carried out: (1) comparing the stress and control groups, and (2) comparing high and low performers. Prior to conducting these group comparisons, the data were examined in order to determine which statistical test would be appropriate. As much of the data turned out to not have normal distributions, the following non-parametric tests were used as appropriate according to the data: the Kruskal-Wallis Test and 2 Sample Wilcoxon Test. Both were implemented in R. Note that the boxplots produced in this thesis were created using the ggplot 2 package in R (Wickham, 2009). Data sets that included outliers more than two standard deviations from the mean were analyzed both with and without the outliers to ensure the results were not being affected significantly by the presence of outliers. A two-tailed alpha level of 0.05 was used for all tests. Because this thesis is exploratory, a large number of comparisons were made. Thus, the overall probability of a type I error is greater than 0.05.

B. COMPARISON OF STRESS AND CONTROL GROUPS ACCORDING TO EYE TRACKING AND MAKE GOAL MEASURES.

1. Comparison of Self-Reported Stress Levels between the Stress and Control Groups

In the post-experiment surveys administered to each subject, the subjects were asked to report the level of stress they experienced during the Make Goal Game in terms of how stressful they found the game in general, the timer, and the experimenter's behavior (induced experimenter stress). Using the 2 Sample Wilcoxon Test, the only significant difference found was between the rank means of the levels of induced experimenter stress. The results are listed in Table 4 in Appendix B. A visual comparison of the stress and control groups by type of self-reported stress is portrayed in Figure 5. In addition to the significant difference between the two groups with regard to experimenter induced stress, there appears to be a trend for the stress group to report higher stress levels in all three stress categories, despite several outliers. It should be noted that in Figure 5, the black lines running through the middle of the boxplots (but seen at the very bottom of the induced stress plots for both the stress and control groups) represent the median stress value. The boxes themselves measure the distance from the lower quartile values to the upper quartile values, as measured by distance from the median value. The "whiskers" represent values 1.5 times the interquartile range as measured from the median. The individual dots represent values that fall outside the range indicated by the "whiskers." Regarding the control group game stress plot, the dot indicating a stress value of "4" represents two subjects reporting a stress level of "4," while the dot at "7" indicates that one individual subject's response was a "7." For the control group induced stress plot, the dot at "2" represents the fact that two control group subjects found the experimenter's presence during the experiment to be stressful, and the dot at "1" indicates that one subject reported a stress level of "1." In the stress group game stress plot the dot at "7" represents one subject's response of a stress level of "7."

Game stress refers to the overall game stress, induced stress refers to induced experimenter stress, and timer stress refers to stress caused by the presence of a timer on the game screen. The specific questions used to elicit these self-reported stress levels can be found in Appendix C.

Figure 5. Self-Reported Stress by Experimental Group on an 8-Point Likert Scale

2. Comparison of Stress and Control Groups According to Eye-tracking Measures

Using the Kruskal-Wallis test, there were no significant differences found between the rank means of the stress and control groups' general eye-tracking measures, but there was a trend for the control group to have a greater number of fixations and fixation shifts. This trend is depicted in Figure 6. A tabular comparison is given in Table 5 in Appendix B.

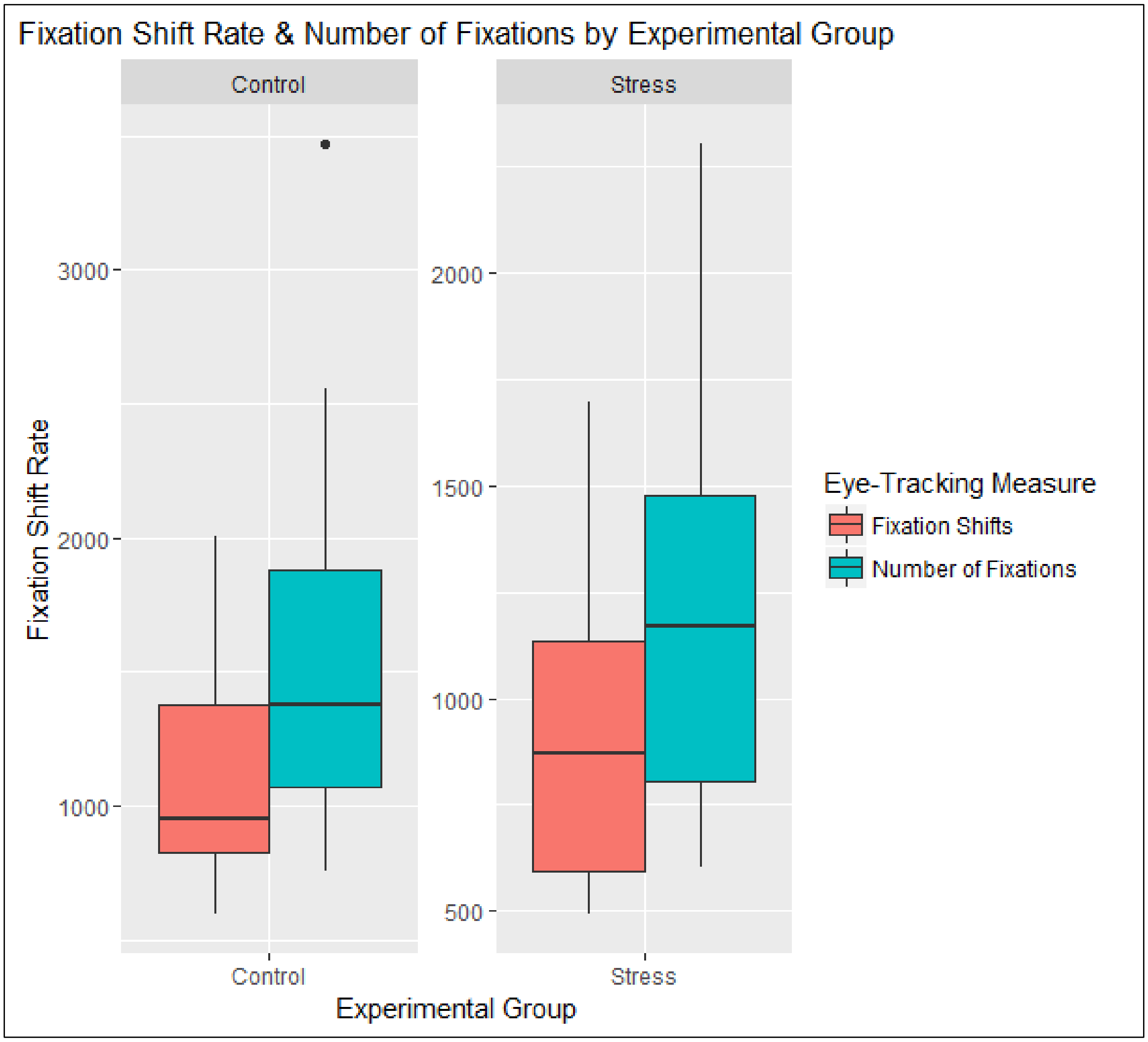


Figure 6. Comparison of Fixations and Fixation Shifts by Experimental Group

The control group's trend towards more fixations is depicted graphically in Figure 7 as a function of game duration. The stress group's fixations over time are also depicted to provide a visual comparison. There appears to be a linear relationship between the control group's fixations versus game duration, implying that the longer the control subjects took to complete the game the more fixations they demonstrated, whereas the stress group displays less of a linear relationship between the two variables.

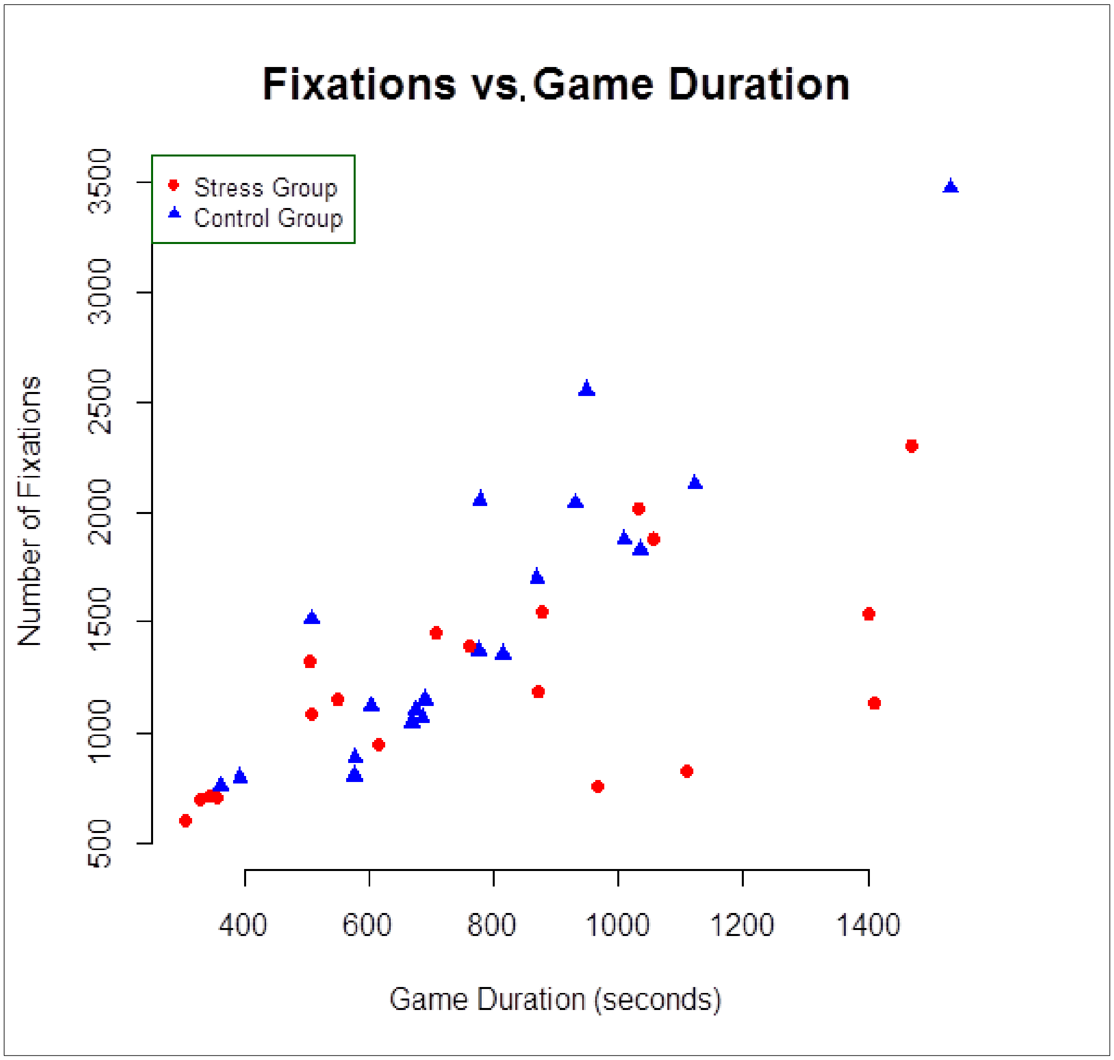


Figure 7. Number of Fixations by Game Duration

There was only one significant difference in the mean percentage of time spent attending to each ROI (as measured by the Kruskal-Wallis Test using rank means). The control group tended to look at the Turns Remaining/Progress ROI more frequently than did the stress group. A tabular summary of the analysis is provided in Table 6, Appendix B. There were outliers in the stress group for the percentage of time spent attending to the Missed Goal and Timer/Stations ROIs. These outliers were removed and the statistical tests were repeated, but the results were the same as those reported in Table 6. Boxplots of the percentage of time spent fixating on each ROI are shown in Figure 8.

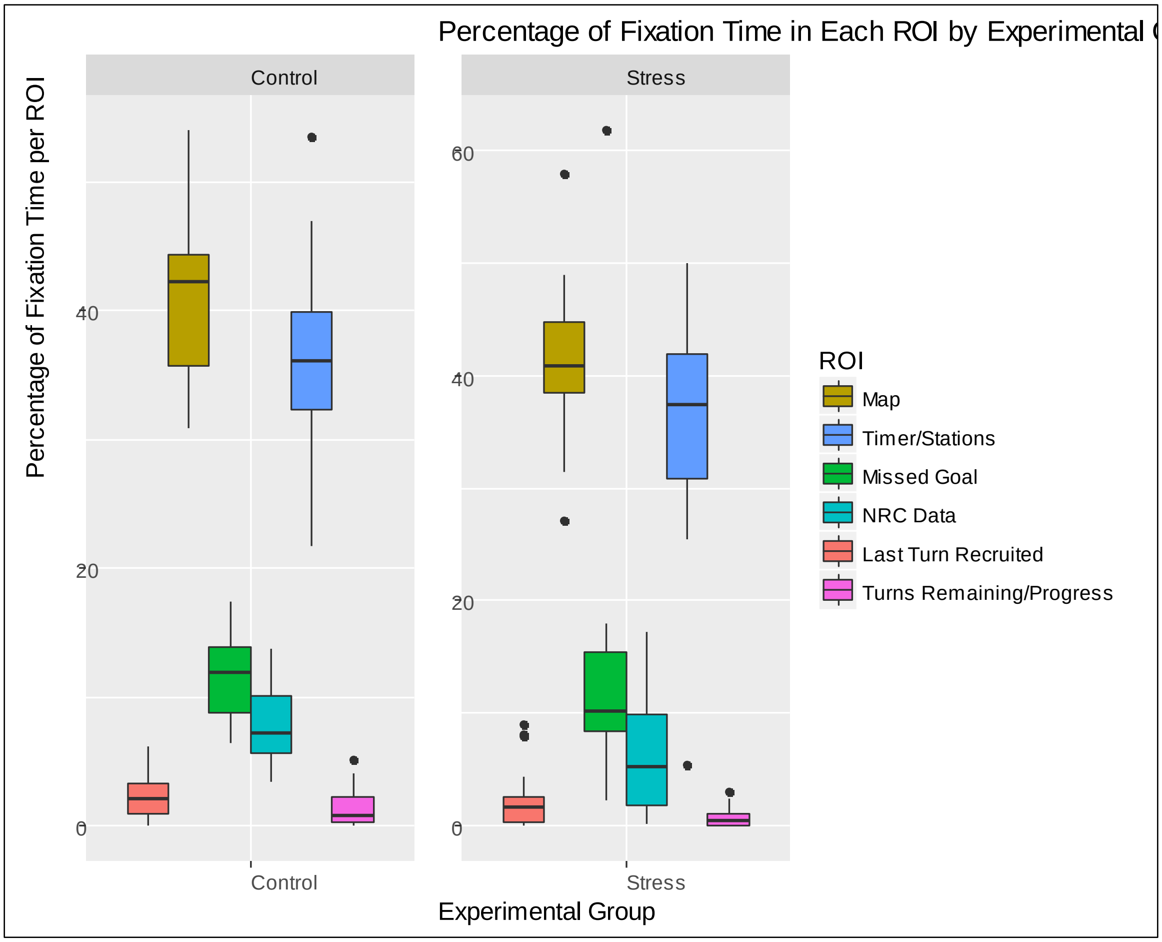


Figure 8. Stress and Control Groups Compared According to Percentage of Fixation Time per ROI.

Using the Kruskal-Wallis test, there were no significant differences found between the stress and control groups in terms of the rank means of duration (in seconds) that subjects in each group looked at the various ROIs. The comparison between the stress and control groups is depicted graphically in Figure 9 and in tabular form in Table 7, Appendix B.

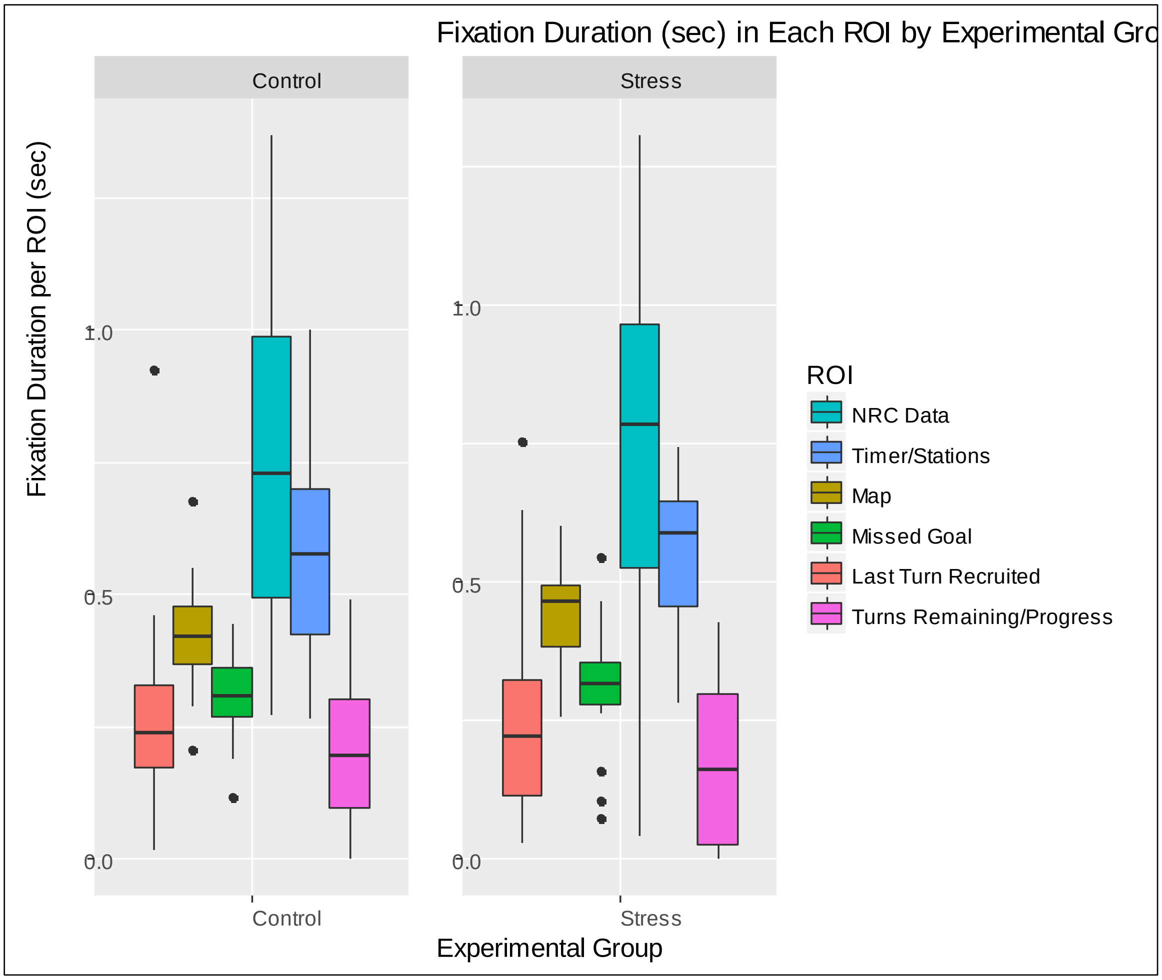


Figure 9. Stress and Control Groups Compared According to Fixation Duration in Seconds per ROI.

3. Comparison of Stress and Control Groups According to Make Goal Performance Measures

Figure 10 visually summarizes the results comparing the stress and control groups according to the Make Goal game outcome measures (see Table 2 in Chapter II for the definition of each measure). A tabular summary can be found in Table 8, Appendix B; no significant differences were found (using the Kruskal-Wallis test) between the rank means of the stress and control groups in terms of any of the Make Goal measures. An outlier was found in the control group for the Accessions variable and the analysis was redone, with no change in results.

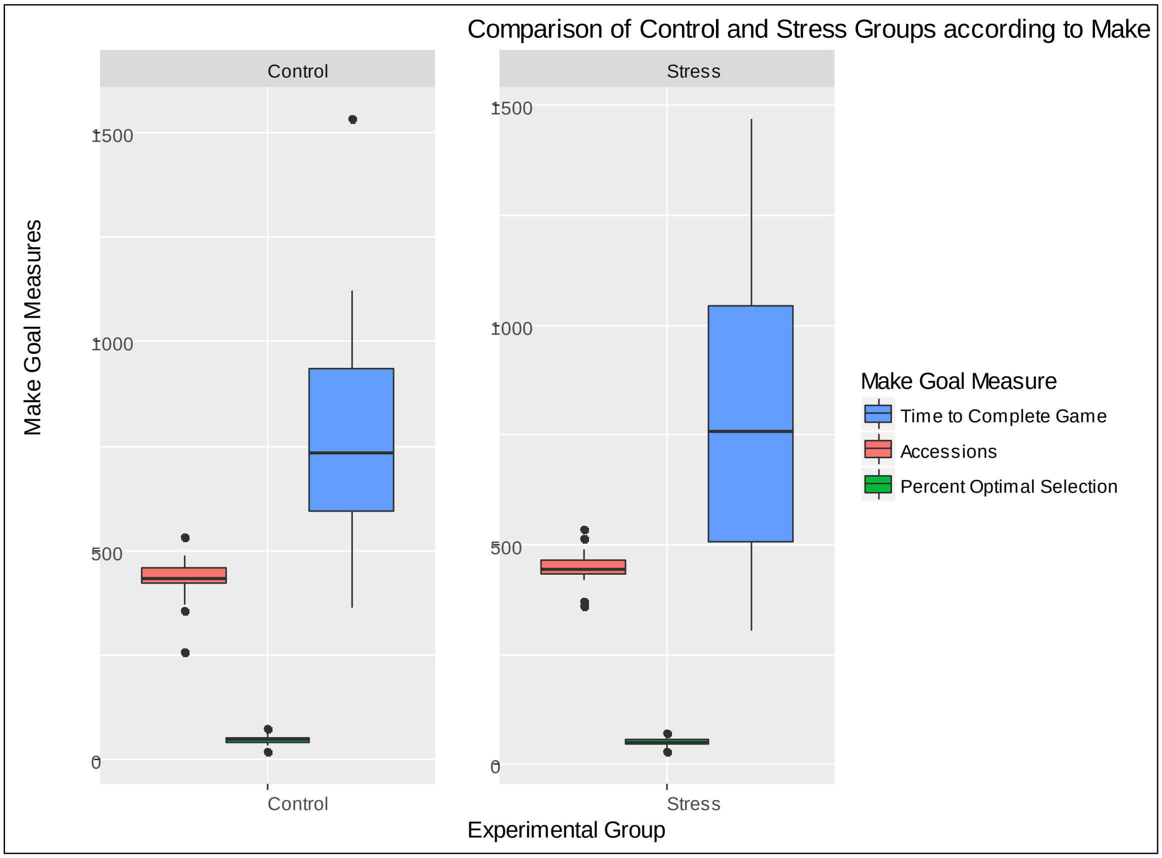


Figure 10. Experimental Group Comparison According to Make Goal Measures

C. COMPARISON OF HIGH AND LOW PERFORMERS ACCORDING TO EYE-TRACKING AND MAKE GOAL MEASURES

The next group comparison made was between the high and low Make Goal performers. High performers were those subjects with the top ten accession scores; low performers were those with the lowest ten accession scores. This method was used because it yields a large enough sample size (n = 20) to be of use statistically. As demonstrated in Table 9, Appendix B, the Kruskal-Wallis test indicates that the high performers did perform significantly better in terms of the rank mean of total accessions and percent optimal station selection than did the low performers, despite taking similar amounts of time to complete the game. The next step was to determine if eye tracking could provide insights into why the high performers did so much better in terms of total accessions gained than did the low performers. A visual comparison between the high and low performing groups is given in Figure 11.

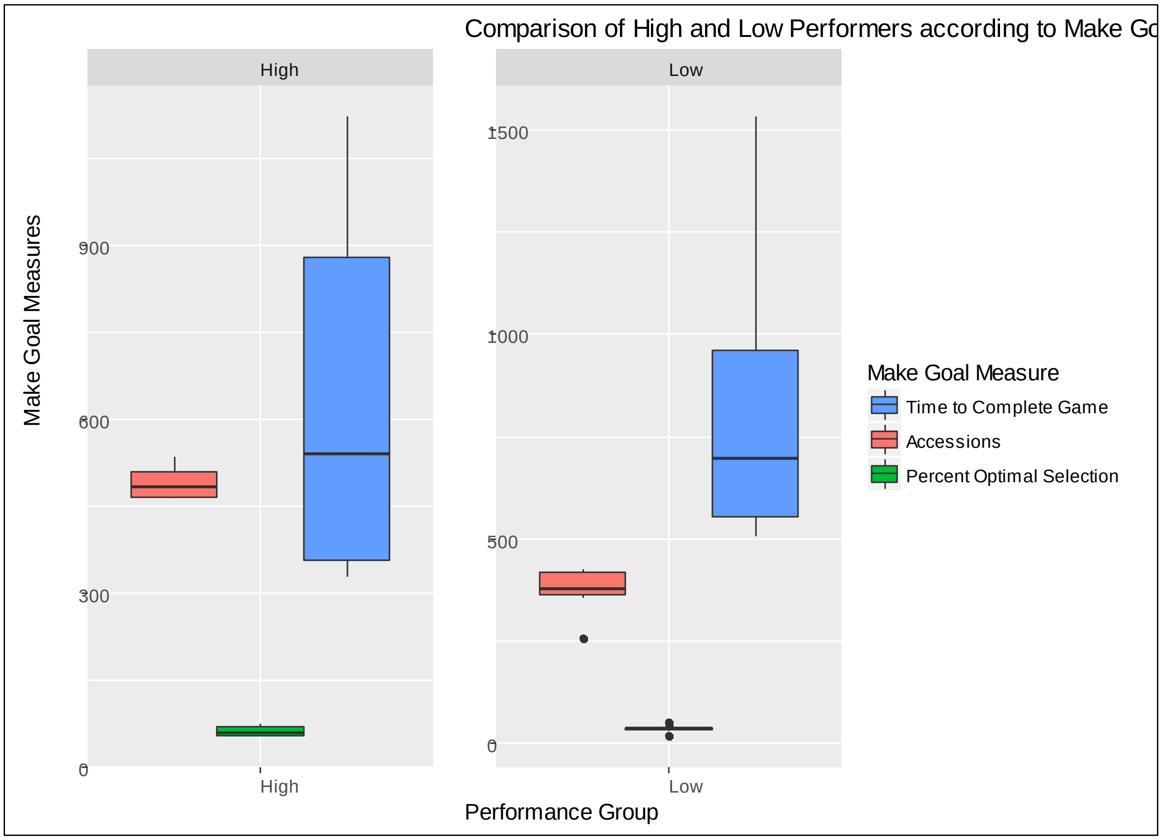


Figure 11. High and Low Performers Compared According to Make Goal Measures

1. Comparison of High and Low Performers According to Eye-Tracking Measures

Using the Kruskal-Wallis test, no significant differences were found between the top and bottom ten performers according the rank means of the general eye-tracking measures. Graphically the groups are compared in Figures 12 and 13 according to these measures (separate graphs were used in order to compare measures on similar scales). A tabular summary can be found in Table 10, Appendix B. It is interesting to note that the top ten performers included five members of the stress group and five members of the control group, while the bottom ten performers included four subjects from the stress group and six from the control group—a fairly even split.

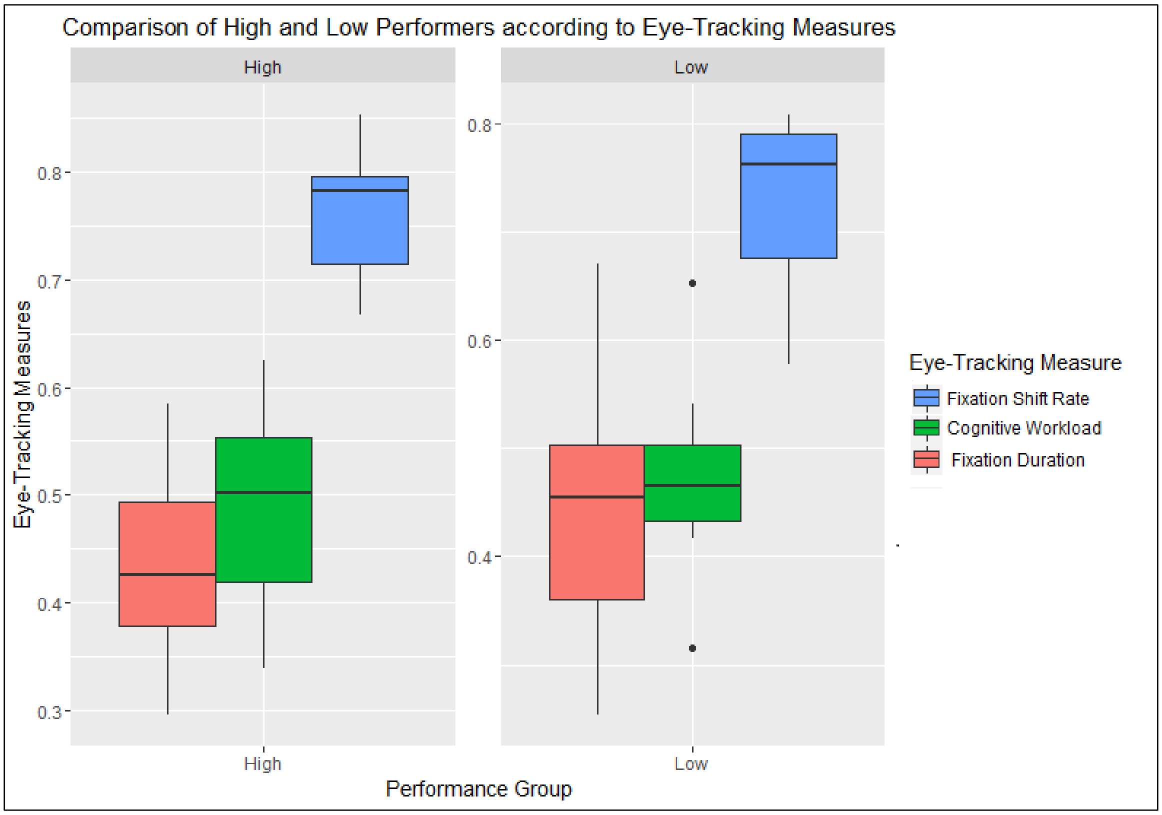


Figure 12. Comparison of High and Low Performers According to Fixation Duration, Cognitive Workload and Fixation Shift Rate

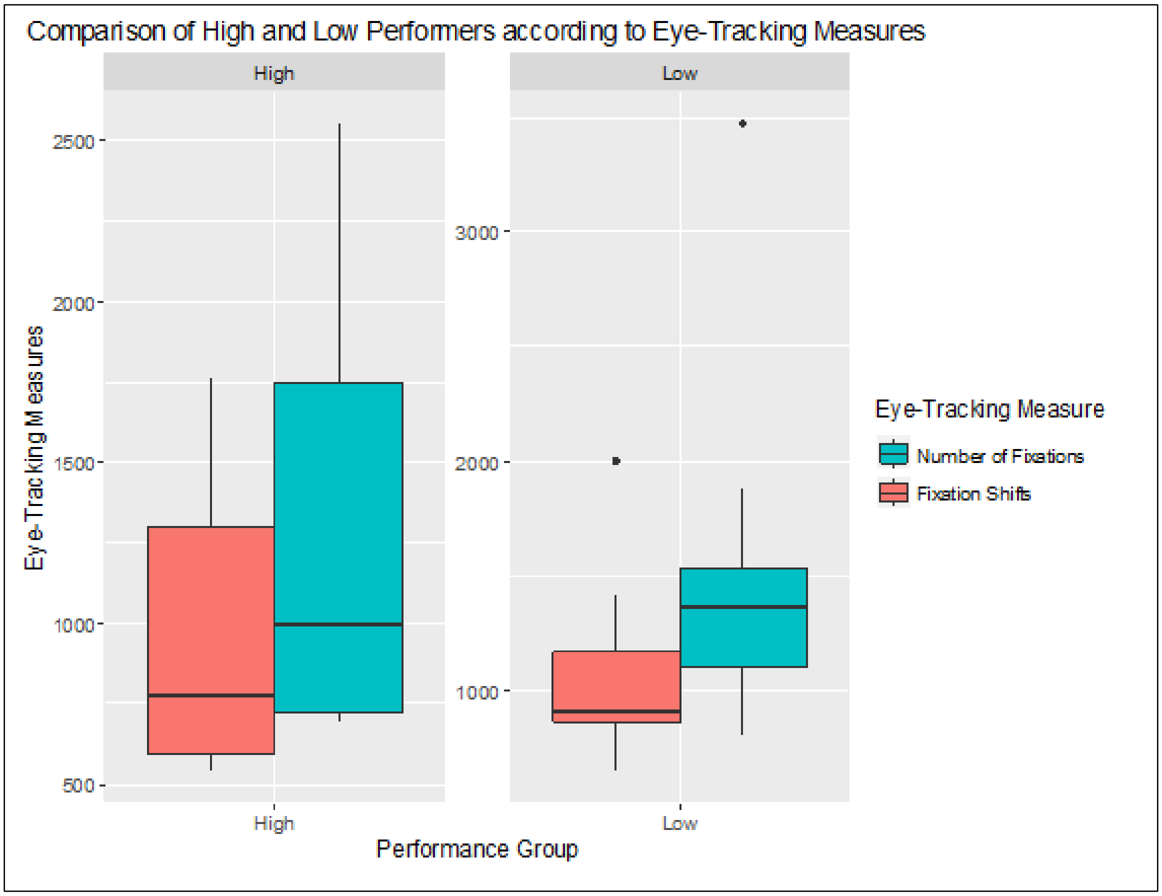


Figure 13. Comparison of High and Low Performers According to Fixation Shifts and Number of Fixations

Of note in Figure 13 is the greater variability seen in the high-performing group's fixations and fixation shifts.

Depicted in Figure 14 are the comparisons between the high and low performers according to the percentage of time spent looking at each ROI. The Kruskal-Wallis test indicates that differences between the two groups are significant for the rank means of percentage of time spent fixating on the Missed Goal ROI and Turns Remaining/Progress ROI. The high-performing group tended to look, on average, more frequently at the Missed Goal ROI and the lower-performing group tended to look, on average, more frequently at the Turns Remaining/Progress ROI. A tabular comparison is given in Table 11, Appendix B.

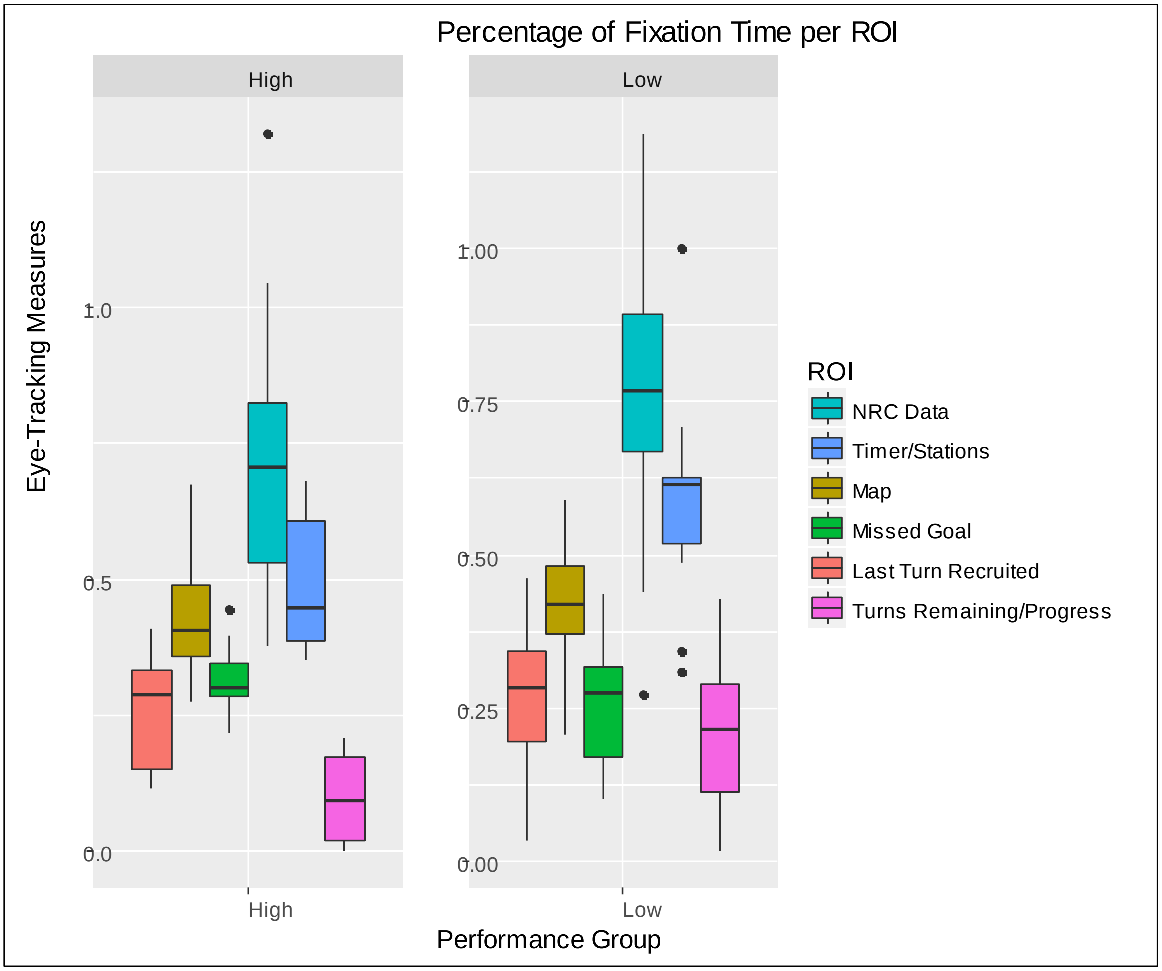


Figure 14. Comparison of High and Low Performers According to Percentage of Fixation Time in Each ROI

Figure 15 depicts the comparison of the top and bottom ten performers according to the amount of time in seconds that they spent looking at each ROI. There were no significant differences found in the rank means of fixation duration using the Kruskal-Wallis test. Table 12 in Appendix B depicts these comparisons in tabular form.

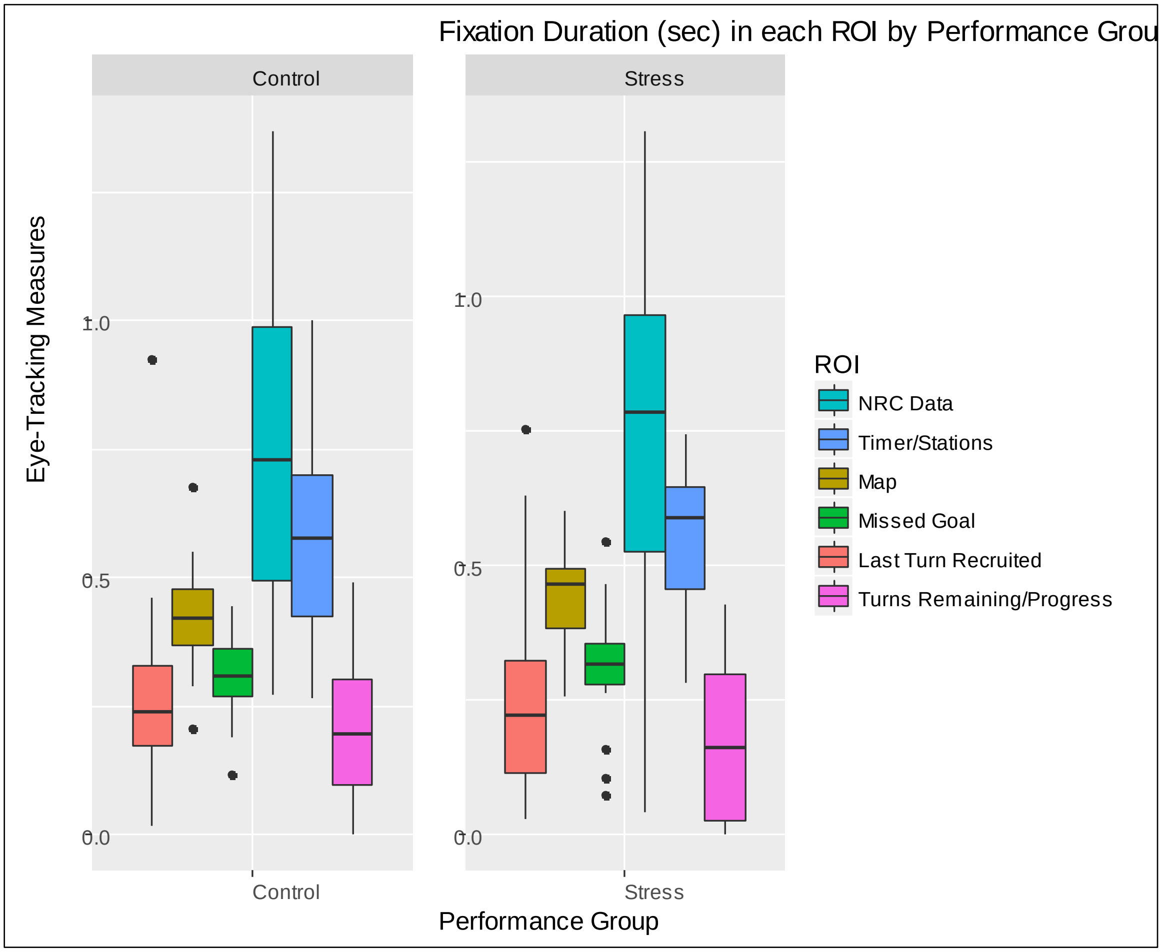


Figure 15. High and Low Performers Compared According to Fixation Duration in Seconds in Each ROI.

Figure 16 depicts results of the analysis of the self-reported stress of the high and low performers. The 2 Sample Wilcoxon test indicates a significant difference in the rank sums of timer stress between the high and low performers, with the high performers reporting a significantly higher level of timer stress than the low performers. Note that there were only two low-performing subjects who reported any timer stress, as indicated by the dots at the stress values of “2” and “4.” A tabular summary of the comparison can be found in Table 13 in Appendix B.

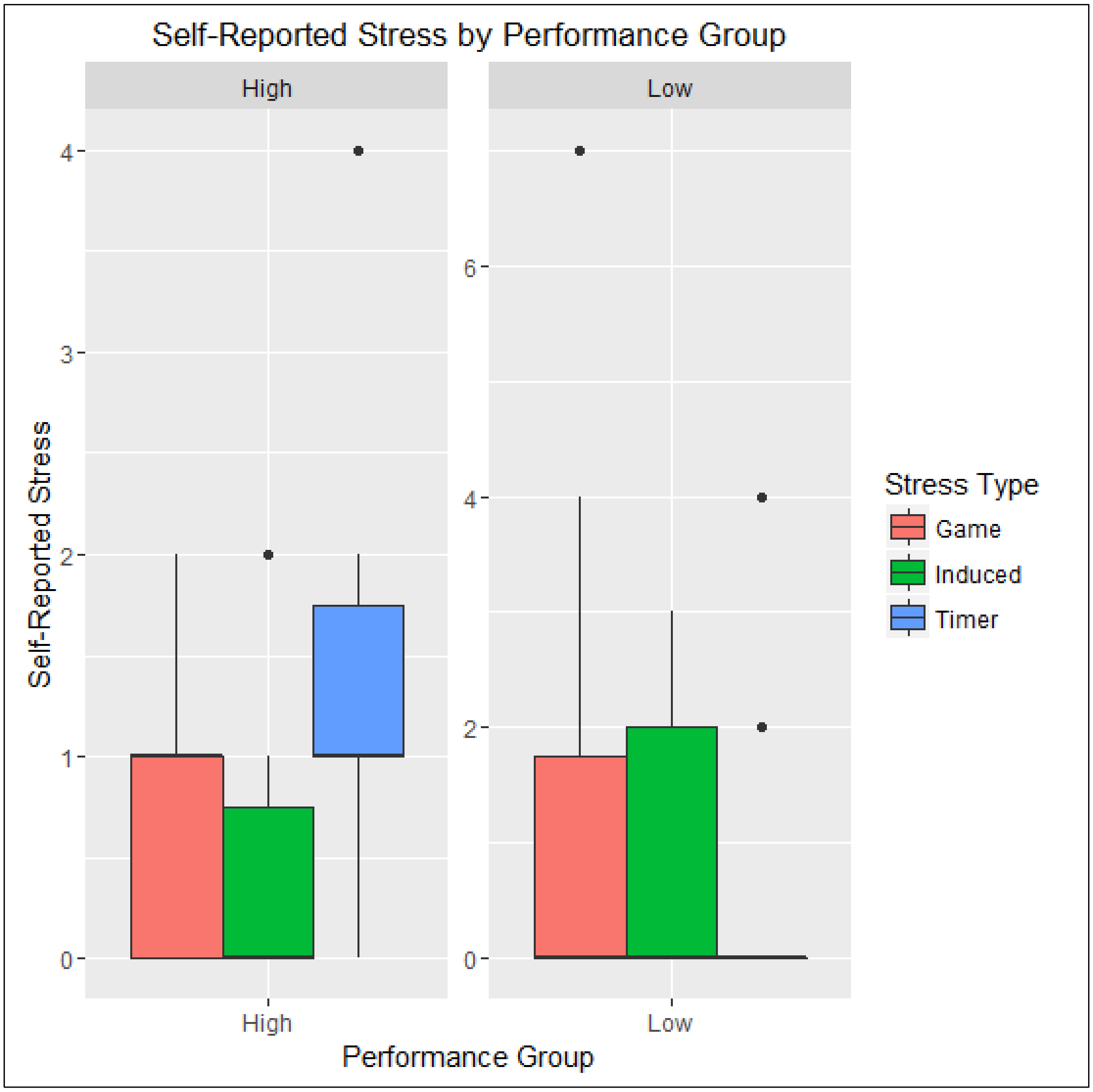


Figure 16. High and Low Performance Groups Compared According to Self-Reported Stress.

2. Summary of Results

In conclusion, the comparison between the stress and control groups yielded few significant differences in rank means or sums, as measured by the Kruskal-Wallis and 2 Sample Wilcoxon tests, respectively. The only difference found was that the control group spent a larger percentage of time, on average, looking at the Turns Remaining/ Progress ROI than did the stress group. The comparison between the high and low performers yielded more significant differences, however. The high performers, as would be expected, had a higher mean number of accessions and a higher mean incidence of percent optimal station selection. High performers were also found, on average, to spend a greater percentage of time looking at the missed goal ROI and the low performers spent more time, on average, looking at the Turns Remaining/Progress ROI.

IV. DISCUSSION

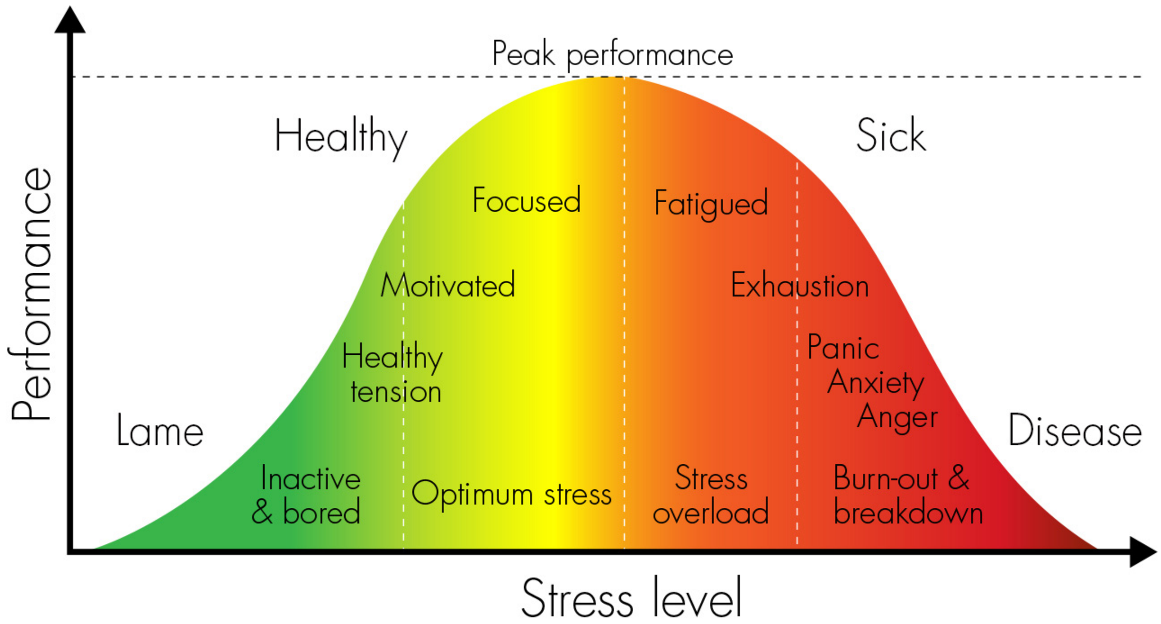
A. INTERPRETATION

1. Overall Impact of Stress

Interpreting the results of this analysis to decipher exactly how stress impacted performance is complicated, because the control group also reported experiencing some level of stress while playing Make Goal, whether from the game itself, the timer, or having the experimenter present while they performed the task (induced experimenter stress), a point discussed below. Additionally, almost no significant differences were found in attention-allocation patterns and workload from the eye-tracking data. Furthermore, there was no statistically significant difference in performance between the stress and control groups playing Make Goal, as depicted in Table 8 (Appendix B, p. 49).

According to the Yerkes-Dodson stress model (Yerkes & Dodson, 1908), moderate levels of stress will result in optimal performance, while very low or very high levels of stress result in degraded performance. It is possible that such a non-monotonic relationship between stress and performance is being exemplified in these results; for example, it is possible that on average, both stress and control subjects experienced low levels of stress, which resulted in similarly modest Make Goal performance and similar attention allocation patterns. However, the experiment would require a greater range of stress levels in order to correlate specific levels of stress with specific points along the Yerkes-Dodson curve, as shown in Figure 17. This experiment only contained two levels of stress (timer only, timer plus social stressor) which cover only the low to moderate stress level range in Figure 17. Thus, an experiment that also includes a no stress level and a high stress level may be more likely to detect stress effects on decision performance and attention allocation patterns.

Interestingly, the high performers reported the timer as significantly more stressful than the low performers. It may be the case that this pushed them to moderate levels of stress, aiding their performance. In contrast, the low performers appear to have experienced low levels of stress, hurting their performance.



Source: Pieterse, Graydon, n.d., *What's Stress Got to Do with It?*, http://haleo.co.uk/wordpress/whats-stress-got-to-do-with-it/

Figure 17. Yerkes-Dodson Inverted U Relationship between Performance and Stress

2. Comparisons between the Stress Group and Control Group on Eye Tracking and Game Performance

As discussed in the Results section, the results showed that there was only one parameter significantly different between the stress and control groups, with the control group looking, on average, more frequently at the Turns Remaining/Progress ROI. There also appeared to be a trend for the stress group to tend to have more fixations than the control group. It is unclear why exactly the control group fixated more on the Turns Remaining/Progress ROI, as it did not provide any important decision making information. A possible explanation is that it is natural in a timed process to want to keep track of time left, or that the control subjects had lost interest in the game and kept checking to see when it would be over, or again, they may have been frustrated trying to figure the game out and were checking how many more turns they had to do so. Control subjects may have fixated on the ROI out of frustration as well; the progress bar displays an accessions goal of seven hundred twenty-five recruits, but the game was designed so that the maximum possible number of accessions was six hundred five. Subjects may have been stressed/panicked looking at how far their progress always remained from the displayed goal.

The results were not quite what was predicted; based on the literature and research regarding stress and its effects on cognitive processes as measured by eye tracking, more significant differences between the stress and control groups would have been expected. The stress group was expected to perform better than the control group, because the stress was assessed to be moderate and in the range likely to enhance performance (according to the Yerkes-Dodson stress model.) However, this was not the case. In addition, higher workload and more fixations would have been expected for the stress group. There was a trend for the stress group to have more fixations, which makes sense, even if it was not statistically significant. It is possible that the sample sizes (n =20) for each group were simply not large enough to capture these expected differences, or that the sample group consisting solely of military officers was less likely to be affected by the experimental stress induction than an equivalent group from a civilian university. For instance, the mild to moderate stress induced for the stress group may not have been enough to faze these specific subjects. Generally speaking, military personnel have at some point in their careers and especially in basic training been exposed to stressors such as being yelled at, told their performance is miserable, etc. Thus, the induced experimental stressor may have been something they were easily able to ignore.

3. Comparison between High and Low Performers on Eye tracking and Game Performance

The comparison between the high- and low-performing groups showed more significant differences. It is interesting to note that the top ten scores were attained by the same number of stress group subjects as control group subjects, and the bottom ten performers included six members of the control group and four members of the stress group. This reinforces the finding that there was no significant difference in performance between the stress and control groups. As would be expected, the high performers had more accessions, and a higher incidence of optimal station selection. The eye-tracking information helped underscore the differences in attention allocation between the two groups. The differences found between the two groups were statistically significant for the mean percentage of total fixations on the Missed Goal ROI and the mean percentage of total fixations on the Turns Remaining/Progress ROI, with the high performers looking more frequently at the Missed Goal ROI and the low performers looking more often at the Turns Remaining/Progress ROI. This is in keeping with what was expected: the Missed Goal ROI provided the most useful feedback to subjects regarding the extent to which each of their decisions deviated from the optimal decision. Turns Remaining/ Progress provided no useful strategic feedback to subjects, so it is not surprising that the subjects who spent more time looking at that ROI were underperforming.

B. IMPLICATIONS

This was an exploratory study and as such had no formal hypotheses; it was hoped that the cognitive processes of better decision makers could be quantified by the eye-tracking data collected, and that the effects of stress on cognitive processes could similarly be quantified according to the eye-tracking data. This study showed that higher performers were in fact better able to figure out and exploit patterns in the game, by attending to the most salient pieces of information. The higher performers also reported experiencing greater stress due to the timer in the game, something that reinforces the concept of the Yerkes-Dodson (1908) model of a non-monotonic relationship between stress and performance. Overall, this study has implications for the Army's goal of quantifying cognitive agility in personnel. Eye-tracking information is a measure that was shown to correlate to performance and provide some insight into what information high performers chose to focus on versus the low performers.

C. LESSONS LEARNED

Several important lessons were learned from this study, particularly with regard to experimental design. Due to the way the game was designed, it was impossible to remove the timer feature for the control subjects and keep it for the stress group. Therefore, the control group was not "stress-free." In retrospect it is not that surprising that the stress and control groups were similar in terms of the stress the subjects experienced. This pattern made interpreting the subjects' performance results in terms of stress much more complicated because both groups reported finding the timer stressful, and even to some extent having an experimenter present, even without stress induction. A future experiment could benefit from this lesson and design an experiment in which the control group receives a more stress-free condition. This could possibly be accomplished by removing the time limit per turn (as the timer was reported as the most stressful aspect of the game by the control group), and having the experimenter leave the room. This way the stress results would not be confounded by having similar stressors shared by both groups. In order to facilitate replication of the Yerkes-Dodson curve over a quantified range of stressors, it would be helpful to design the experiment to include a means of objectively quantifying stress on a continuous scale, perhaps using an EEG, heart rate monitor, or cortisol levels. This way stress levels could be clearly quantified and plotted against performance, which would be very informative.

In trying to interpret the results of the eye-tracking data in the various ROIs, it was found that the control group and low performers both tended to look at the Turns Remaining/Progress ROI more frequently than the other ROIs. As was discussed above, there are several possible interpretations for this, and it would be interesting to know what the individual subject's motivations and logic were in focusing on this ROI. A future study could add a question to the Post Task Survey asking subjects to explain why they looked in each region to help clarify what their cognitive process was.

D. CONCLUSION

The ultimate goal of the overall TRADOC study was first, to identify cognitively agile military personnel and second, to investigate the effects of stress on cognitive agility. In support of these goals the aim of this thesis was to 1) determine if certain attention-allocation patterns are associated with effective cognitive-agility performance, and 2) to investigate the effects of stress on cognitive agility as measured by attention allocation. The study was a success in that the eye-tracking data did, in fact, provide insight into what information the higher-performing (more cognitively agile) subjects attended to, revealing common attention-allocation patterns among high performers versus low performers. In terms of the effects of stress on cognitive agility, the precise effects in this study were difficult to quantify due to experimental design issues, but important lessons were learned that will hopefully be of use in designing and executing future studies.

From the analysis of the eye-tracking data collected on the forty subjects who played the Make Goal game, it became evident that there was only one significant difference between the stress and control groups. This was an unexpected difference, which involved the control group tending to fixate more frequently on an ROI containing non-essential information (Turns Remaining/Progress ROI). High performers tended to spend a greater percentage of time looking at the missed goal ROI (the most essential information) while the low performers tended to spend more time looking at the Turns Remaining/Progress ROI. In this regard, the eye tracking did provide insight into what the high performers chose to focus on versus the low performers, which shows promise for future cognitive-agility studies.

The effects of stress were not as clear from the analysis, and interestingly, the initial expectation that the stress group would perform differently was not borne out by the statistical analysis. This is likely the case because the control group was not entirely stress-free. Based on this finding it is recommended that future studies use a greater number of stress levels and incorporate a means of objectively quantifying stress levels (i.e., via skin conductance, EEG), in order to more effective apply the Yerkes-Dodson paradigm to the analysis. In conclusion, this study obtained some useful results in terms of using eye- tracking data to determine why some subjects performed well while others performed poorly. This study was exploratory, and it is hoped that the lessons learned will pave the way for future study in this area.

APPENDIX A. MAKE GOAL DATA ATTRIBUTES BY TURNS

Figures 18 to 23 depict the six times during the Make Goal game that the optimal station selection and most important attribute changed, as referenced in Section B, part 1, describing the Make Goal task (p. 10–11). The first four columns labeled "Laredo," "McAllen," "Mercado" and "N. Corpus" are all recruiting districts in which subjects could place the recruiter. The numbers given in each column indicate the total number of recruits that could be gained by placing the recruiter in that district for that particular turn. The "Max Goal" column indicates the highest possible number of recruits that could possibly be gained on that turn by selecting the optimal station. The "Time per Turn" column indicates the number of seconds subjects had on that particular turn to place the recruiter in a given district. The attribute column indicates the particular data attribute that was most salient (and therefore distinguished the optimal station) for that particular turn.

Laredo	McAllen	Mercado	N. Corpus	Max Goal	Time per Turn (sec)	Attribute	Optimal Station
5	3	2	0	5	100	Diversity	Laredo
6	4	2	0	6	100	Diversity	Laredo
4	3	1	2	4	100	Diversity	Laredo
6	0	2	1	6	100	Diversity	Laredo
9	7	0	4	9	100	Diversity	Laredo
3	4	0	1	4	100	Diversity	McAllen
5	4	2	3	5	90	Diversity	Laredo
3	3	2	3	3	90	Diversity	Laredo/McAllen/N.Corpus
5	4	1	2	5	90	Diversity	Laredo
8	0	3	2	8	90	Diversity	Laredo
5	6	3	2	6	90	Diversity	McAllen
5	4	0	0	5	90	Diversity	Laredo
8	6	3	3	8	80	Diversity	Laredo
7	4	3	3	7	80	Diversity	Laredo
8	6	1	1	8	80	Diversity	Laredo
8	5	1	1	8	80	Diversity	Laredo
6	4	2	0	6	80	Diversity	Laredo
6	6	3	1	6	80	Diversity	Laredo/McAllen
10	4	1	1	10	70	Diversity	Laredo
4	4	1	0	4	70	Diversity	Laredo/McAllen

Figure 18. Optimal Attribute/Recruiting Station Selections for Moves 1–20 of Make Goal

Laredo	McAllen	Mercado	N. Corpus	Max Goal	Time per Turn (sec)	Attribute	Optimal Station
0	3	5	2	5	70	QMA	Mercado
0	4	6	2	6	70	QMA	Mercado
2	3	4	1	4	70	QMA	Mercado
1	0	6	2	6	70	QMA	Mercado
4	7	9	0	9	60	QMA	Mercado
1	4	3	0	4	60	QMA	McAllen
3	4	5	2	5	60	QMA	Mercado
3	3	3	2	3	60	QMA	Mercado/Laredo/McAllen
2	4	5	1	5	60	QMA	Mercado
2	0	8	3	8	60	QMA	Mercado
2	6	5	3	6	50	QMA	McAllen
0	4	5	0	5	50	QMA	Mercado
3	6	8	3	8	50	QMA	Mercado
3	4	7	3	7	50	QMA	Mercado
1	6	8	1	8	50	QMA	Mercado
1	5	8	1	8	50	QMA	Mercado
0	4	6	2	6	40	QMA	Mercado
1	6	6	3	6	40	QMA	Mercado/McAllen

Figure 19. Optimal Attribute/Recruiting Station Selections for Moves 21–38 of Make Goal

Laredo	McAllen	Mercado	N. Corpus	Max Goal	Time per Turn (sec)	Attribute	Optimal Station
5	3	2	0	5	40	Diversity	Laredo
6	4	2	0	6	40	Diversity	Laredo
4	3	1	2	4	40	Diversity	Laredo
6	0	2	1	6	40	Diversity	Laredo
9	7	0	4	9	30	Diversity	Laredo
3	4	0	1	4	30	Diversity	McAllen
5	4	2	3	5	30	Diversity	Laredo
3	3	2	3	3	30	Diversity	Laredo/McAllen/N.Corpus
5	4	1	2	5	30	Diversity	Laredo
8	0	3	2	8	30	Diversity	Laredo
5	6	3	2	6	25	Diversity	McAllen
5	4	0	0	5	25	Diversity	Laredo
8	6	3	3	8	25	Diversity	Laredo
7	4	3	3	7	25	Diversity	Laredo
8	6	1	1	8	25	Diversity	Laredo
8	5	1	1	8	25	Diversity	Laredo
6	4	2	0	6	20	Diversity	Laredo

Figure 20. Optimal Attribute/Recruiting Station Selections for Moves 39–55 of Make Goal

Laredo	McAllen	Mercado	N. Corpus	Max Goal	Time per Turn (sec)	Attribute	Optimal Station
2	5	3	0	5	20	YBC	McAllen
2	6	4	0	6	20	YBC	McAllen
1	4	3	2	4	20	YBC	McAllen
2	6	0	1	6	20	YBC	McAllen
0	9	7	4	9	20	YBC	McAllen
0	3	4	1	4	18	YBC	Mercado
2	5	4	3	5	18	YBC	McAllen
2	3	3	3	3	18	YBC	McAllen/Mercado/N.Corpus
1	5	4	2	5	18	YBC	McAllen
3	8	0	2	8	18	YBC	McAllen
3	5	6	2	6	18	YBC	Mercado
0	5	4	0	5	17	YBC	McAllen
3	8	6	3	8	17	YBC	McAllen
3	7	4	3	7	17	YBC	McAllen
1	8	6	1	8	17	YBC	McAllen
1	8	5	1	8	17	YBC	McAllen

Figure 21. Optimal Attribute/Recruiting Station Selections for Moves 56–71 of Make Goal

Laredo	McAllen	Mercado	N. Corpus	Max Goal	Time per Turn (sec)	Attribute	Optimal Station
0	3	5	2	5	17	QMA	Mercado
0	4	6	2	6	16	QMA	Mercado
2	3	4	1	4	16	QMA	Mercado
1	0	6	2	6	16	QMA	Mercado
4	7	9	0	9	16	QMA	Mercado
1	4	3	0	4	16	QMA	McAllen
3	4	5	2	5	16	QMA	Mercado
3	3	3	2	3	15	QMA	Mercado/McAllen/Laredo
2	4	5	1	5	15	QMA	Mercado
2	0	8	3	8	15	QMA	Mercado
2	6	5	3	6	15	QMA	McAllen
0	4	5	0	5	15	QMA	Mercado
3	6	8	3	8	15	QMA	Mercado
3	4	7	3	7	15	QMA	Mercado
0	4	6	2	6	15	QMA	Mercado
2	3	4	1	4	15	QMA	Mercado

Figure 22. Optimal Attribute/Station Selections for Moves 72–87 of Make Goal

Laredo	McAllen	Mercado	N. Corpus	Max Goal	Time per Turn (sec)	Attribute	Optimal Station
9	7	0	4	9	15	Diversity	Laredo
3	4	0	1	4	15	Diversity	Laredo
5	4	2	3	5	15	Diversity	Laredo
3	3	2	3	3	15	Diversity	Laredo/McAllen/N.Corpus
5	4	1	2	5	15	Diversity	Laredo
8	0	3	2	8	15	Diversity	Laredo
5	6	3	2	6	15	Diversity	McAllen
5	4	0	0	5	15	Diversity	Laredo
8	6	3	3	8	15	Diversity	Laredo
7	4	3	3	7	15	Diversity	Laredo
8	6	1	1	8	15	Diversity	Laredo
8	5	1	1	8	15	Diversity	Laredo
6	4	2	0	6	15	Diversity	Laredo

Figure 23. Optimal Attribute/Recruiting Station Selections for Moves 88–100 of Make Goal

APPENDIX B. DATA ANALYSIS SUMMARY TABLES

Table 4. Subject Self-Reported Stress

Stressor	Control (mean,sd) 95% CI	Stress (mean,sd) 95% CI	Statistical Test	Result
Game	1.2 (1.85) [0.33, 2.07]	1.42 (1.92) [0.49, 2.35]	2 Sample Wilcoxon	W= 177 p-value = 0.71 No significant difference
Timer	1.1 (1.33) [0.48, 1.72]	1.95 (2.09) [0.94, 2.96]	2 Sample Wilcoxon	W = 151.5 p-value= 0.26 No significant difference
Induced Stress	0.25 (0.64) [0, 0.55]	1.05 (1.27) [0.44, 1.66]	2 Sample Wilcoxon	W = 123 p-value = 0.02 Significant difference

Table 5. Comparison of Control and Stress Groups According to General Eye-Tracking Measures

Measure	Control (mean,sd) 95% CI	Stress (mean,sd) 95% CI	Statistical Test	Result
Overall Fixation Duration	0.45 (0.11) [0.39,0.50]	0.46 (0.09) [0.42,0.50]	Kruskal-Wallis	χ^2 (1) = 0.10 p-value = 0.18 No significant difference
Number of Fixations	1534 (689) [1211,1856]	1224 (485) [990,1458]	Kruskal-Wallis	χ^2 (1) = 1.82 p-Value= 0.18 No significant difference
Fixation Shifts	1088 (350) 903,1273]	902 (340) [782,1066]	Kruskal-Wallis	χ^2 (1) = 1.82 p-value = 0.38 No significant difference
Fixation Shift Rate	0.73 (0.08) [0.69,0.77]	0.75 (0.07) [0.71,0.78]	Kruskal-Wallis	χ^2 (1) = 0.13 p-value= 0.71 No significant difference
Cognitive Workload	0.51 (0.09) [0.47,0.55]	0.48 (0.10) [0.43,0.53]	Kruskal-Wallis	χ^2 (1) = 0.38 p-value = 0.54 No significant difference

Table 6. Comparison of Control and Stress Groups According to Percentage of Fixation Time in Each ROI

Variable Name ROI	Control mean(sd) 95% CI	Stress mean(sd) 95% CI	Statistical Test	Result
Timer/ Stations	36.36 (7.41) [32.89,39.83]	35.95 (9.98) [31.14,40.76]	Kruskal-Wallis	χ^2 (1) = 0.08 p-value = 0.78 No significant difference
NRC Data	7.81 (3.13) [6.35,9.28]	6.61 (5.26) [4.07,9.14]	Kruskal-Wallis	χ^2 (1) = 1.20 p-value = 0.27 No significant difference
Missed Goal	11.64 (3.33) [10.09,13.20]	13.55 (12.49) [7.53,19.57]	Kruskal-Wallis	χ^2 (1) = 0.007 p-value = 0.93 No significant difference
Map	40.35 (6.39) [37.36,43.34]	40.77 (7.04) [37.38,44.16]	Kruskal-Wallis	χ^2 (1) = 0 p-value = 1 No significant difference
Last Turn Recruited	2.23 (1.68) [1.44,2.23]	2.40 (2.90) [1.00,2.40]	Kruskal-Wallis	χ^2 (1) = 0.42 p-value = 0.52 No significant difference
Turns Remaining/ Progress	1.61 (1.56) [0.88,2.34]	0.73 (0.88) [0.30,1.15]	Kruskal-Wallis	χ^2 (1) = 4.56 p-value = 0.03 Significant difference

Table 7. Comparison of Control and Stress Groups According to Fixation Duration (in Seconds) in Each ROI

Variable Name ROI	Control (mean,sd) 95 % CI	Stress (mean,sd) 95% CI	Statistical Test	Result
Timer/ Stations	0.56 (0.19) [0.47,0.65]	0.56 (0.13) [0.49,0.62]	Kruskal-Wallis	χ^2 (1) = 0.003 p-value = 0.96 No significant difference
NRC Data	0.78 (0.32) [0.63,0.93]	0.74 (0.37) [0.56,0.92]	Kruskal-Wallis	χ^2 (1) = 0.08 p-value = 0.78 No significant difference
Missed Goal	0.30 (0.08) [0.27,0.34]	0.31 (0.11) [0.25,0.36]	Kruskal-Wallis	χ^2 (1) = 0.08 p-value = 0.78 No significant difference
Map	0.42 (0.10) [0.37,0.47]	0.45 (0.10) [0.40,0.49]	Kruskal-Wallis	χ^2 (1) = 0.71 p-value = 0.40 No significant difference
Last Turn Recruited	0.27(0.19) [0.18,0.36]	0.25 (0.20) [0.15,0.34]	Kruskal-Wallis	χ^2 (1) = 0.53 p-value = 0.47 No significant difference
Turns Remaining/ Progress	0.45 (0.11) [0.39,0.50]	0.46(0.09) [0.42,0.50]	Kruskal-Wallis	χ^2 (1) = 0.78 p-value = 0.38 No significant difference

Table 8. Comparison of Control and Stress Groups According to Make Goal Measures

Measure	Control (mean,sd) 95% CI	Stress (mean,sd) 95 % CI	Statistical Test	Result
Accessions	429 (56.26) [402.67,455.33]	447.47 (41.43) [427.50,467.44]	Kruskal-Wallis	χ^2 (1) = 1.11 p-value = 0.30 No significant Difference
Regret	176.0 (56.26) [149.67, 202.33]	157.33 (41.43) [137.56,177.50]	Kruskal-Wallis	χ^2 (1) = 1.95 p-value = 0.16 No significant difference
% Optimal Selection	46.55 (11.60) [41.02,51.88]	51.10 (10.36) [46.11,56.10]	Kruskal-Wallis	χ^2 (1) = 1.98 p-value = 0.16 No significant difference
Time	777.55 (272.79) [649.89,905.22]	798.49 (379.55) [615.55,981.42]	Kruskal-Wallis	χ^2 (1) = 0.01 p-value = 0.93 No significant difference

Table 9. Comparison of High and Low Performers According to Make Goal Measures

Measure	High (mean,sd) 95 % CI	Low (mean,sd) 95 % CI	Statistical Test	Result
Accessions	490 (27.84) [470.09, 509.91]	378.5 (50.30) [342.52, 414.48]	Kruskal-Wallis	χ^2 (1) = 14.32 p-value = 0.00 Significant difference
Regret	115 (27.84) [95.09,134.91]	226.5 (50.29) [190.52,262.48]	Kruskal-Wallis	χ^2 (1) = 14.32 p-value = 0.00 Significant difference
% Optimal Selection	61 (8) [55.28,66.72]	36.6 (9.36) [29.90,43.30]	Kruskal-Wallis	χ^2 (1) = 13.78 p-value = 0.00 Significant difference
Time	627.13 (310.06) [405.32,848.93]	828.96 (373.00) [562.14,1095.79]	Kruskal-Wallis	χ^2 (1) = 1.85 p-value = 0.17 No significant difference

Table 10. Comparison of High and Low Performers According to General Eye-Tracking Measures

Measure	High (mean,sd) 95 % CI	Low (mean,sd) 95 % CI	Statistical Test	Result
Overall Fixation Duration	0.43 (0.09) [0.37,0.49]	0.44 (0.13) [0.35,0.54]	Kruskal-Wallis	χ^2 (1) = 0.09 p-value = 0.76 No significant difference
Number of Fixations	1274.7 (680.66) [787.79,1761.61]	1513.8 (756.63) [972.54,2055.06]	Kruskal-Wallis	χ^2 (1) = 1.29 p-value= 0.26 No significant difference
Fixation Shifts	944.6 (445.77) [625.72,1263.48]	1063.3 (401.35) [776.19,1350.41]	Kruskal-Wallis	χ^2 (1) = 1.29 p-value = 0.26 No significant difference
Fixation Shift Rate	0.76 (0.06) [0.72,0.80]	0.73 (0.08) [0.67,0.79]	Kruskal-Wallis	χ^2 (1) = 0.82 p-value= 0.36 No significant difference
Cognitive Workload	0.49 (0.09) [0.42,0.56]	0.47 (0.09) [0.41,0.54]	Kruskal-Wallis	χ^2 (1) = 0.37 p-value = 0.54 No significant difference

Table 11. Comparison of High and Low Performers According to Percentage of Fixation Time in Each ROI

Variable Name ROI	High mean(sd) 95% CI	Low mean(sd) 95% CI	Statistical Test	Result
Timer/ Stations	36.86 (6.52) [32.19,41.52]	38.36 (7.06) [33.30,43.41]	Kruskal-Wallis	χ^2 (1) = 0.57 p-value = 0.45 No significant difference
NRC Data	5.72 (3.80) [3.00,8.44]	7.16 (4.37) [4.04,10.28]	Kruskal-Wallis	χ^2 (1) = 0.57 p-value = 0.45 No significant difference
Missed Goal	13.87 (3.25) [11.54, 16.19]	8.98 (3.59) [6.41,11.55]	Kruskal-Wallis	χ^2 (1) = 6.61 p-value = 0.01 Significant difference
Map	40.95 (5.73) [36.85,45.05]	40.52 (4.94) [36.99,44.06]	Kruskal-Wallis	χ^2 (1) = 0.37 p-value = 0.54 No significant difference
Last Turn Recruited	2.30 (2.32) [0.64,3.96]	3.45(2.96) [1.33,5.57]	Kruskal-Wallis	χ^2 (1) = 0.97 p-value = 0.33 No significant difference
Turns Remaining/ Progress	0.31 (0.31) [0.08,0.53]	1.54 (1.31) [0.60, 2.48]	Kruskal-Wallis	χ^2 (1) =6.23 p-value = 0.01 Significant difference

Table 12. Comparison of High and Low Performers According to Fixation Duration (in Seconds) in Each ROI

Variable Name ROI	High (mean,sd) 95 % CI	Low (mean,sd) 95% CI	Statistical Test	Result
Timer/ Stations	0.49 (0.13) [0.40,0.59]	0.59 (0.19) [0.46,0.73]	Kruskal-Wallis	χ^2 (1) = 0.97 p-value = 0.32 No significant difference
NRC Data	0.73 (0.29) [0.53,0.93]	0.77 (0.29) [0.56,0.97]	Kruskal-Wallis	χ^2 (1) = 0.28 p-value = 0.60 No significant difference
Missed Goal	0.32 (0.07) [0.27,0.37]	0.26 (0.11) [0.18,0.33]	Kruskal-Wallis	χ^2 (1) = 0.08 p-value = 0.78 No significant difference
Map	0.42 (0.10) [0.37,0.47]	0.45 (0.10) [0.40,0.49]	Kruskal-Wallis	χ^2 (1) = 1.65 p-value = 0.20 No significant difference
Last Turn Recruited	0.26 (0.12) [0.35,0.52]	0.42 (0.13) [0.17,0.35]	Kruskal-Wallis	χ^2 (1) = 0.01 p-value = 0.47 No significant difference
Turns Remaining/ Progress	0.45 (0.11) [0.39,0.50]	0.46(0.09) [0.42,0.50]	Kruskal-Wallis	χ^2 (1) = 0.78 p-value = 0.94 No significant difference

Table 13. Exploratory Analyses: Comparison of Self-Reported Stress Levels between High and Low Performers

Stressor	High (mean,sd) 95% CI	Low (mean,sd) 95% CI	Statistical Test	Result
Game	0.8 (0.79) [0.24,1.36]	1.40 (2.37) [0, 3.09]	2 Sample Wilcoxon	W= 53 p-value = 0.84 No significant difference
Timer	1.40 (1.07) [0.63, 2.17]	0.60 (1.35) [0, 1.57]	2 Sample Wilcoxon	W = 78.5 p-Value= 0.02 Significant difference
Induced Stress	0.40 (0.70) [0, 0.90]	0.90 (1.20) [0.04, 1.76]	2 Sample Wilcoxon	W = 40.5 p-value = 0.42 No significant difference

APPENDIX C. POST TASK SURVEY

Figure 24 depicts the two pages of the Post Task Survey administered to subjects immediately after they completed the Make Goal task.

Postgame Survey

Subject ID:____________ Date:____________

1) What was your general strategy for assigning recruiters?

2) Did your strategy change during the game? If so, how?

3) Which items of information did you find most useful?

4) How challenging did you find the game?

0	1	2	3	4	5	6	7	8
Not at all challenging								Very challenging

b) If you circled a 1 or higher, which aspects of the game were challenging?

5) How stressful did you find the game?

0	1	2	3	4	5	6	7	8
Not at all stressful								Very stressful

b) If you circled a 1 or higher, which aspects of the game were stressful?

NPS IRB
APPROVED APR 22 2016
EXPIRED APR 10 2017

Figure 24. Make Goal Post Task Survey (Continued on Next Page)

6) Towards the end of the game, how stressful did you find the timer?

0	1	2	3	4	5	6	7	8
Not at all stressful								Very stressful

7) How stressful did you find the experimenter's behavior while you played the game?

0	1	2	3	4	5	6	7	8
Not at all stressful								Very stressful

b) If you circled a 1 or higher, which aspects of their behavior were stressful?

8) Any other comments?

NPS IRB
APPROVED APR 22 2016
EXPIRED APR 10 2017

Figure 24 (Continued from Previous Page)

LIST OF REFERENCES

Alt, J., Appleget, J., Buttrey, S., Kennedy, Q., Sciarini, L., & Johnson, E. (2015). Make goal: A training system for naval recruiting district leaders. In *Report to the Navy Recruitment Command.* Naval Postgraduate School, Monterey, CA

Ashcraft, M. H. (2002). Math anxiety: Personal, educational, and cognitive consequences. *Current Directions in Psychological Science*, *11*(5), 181–185. doi:10.1111/1467-8721.00196

Ashcraft, M. H., & Kirk, E. P. (2001). The relationships among working memory, math anxiety, and performance. *Journal of Experimental Psychology: General*, *130*(2), 224–237. doi:10.1037/0096-3445.130.2.224

Bellenkes, A. H., Wickens, C. D., & Kramer, A. F. (1997). Visual scanning and pilot expertise: The role of attentional flexibility and mental model development. *Aviation, Space, and Environmental Medicine, 68*(7), 569–579. Retrieved from https://www.researchgate.net/publication/14004960_Visual_scanning_and_pilot_expertise_The_role_of_attentional_flexibility_and_mental_model_development

Braunstein-Bercovitz, H., Dimentman-Ashkenazi, I., & Lubow, R. E. (2001). Stress affects the selection of relevant from irrelevant stimuli. *Emotion 1*(2), 182–192.

Braunstein-bercovitz, H. (2003). Does stress enhance or impair selective attention? The effects of stress and perceptual load on negative priming. *Anxiety, Stress, & Coping*, *16*(4), 345–357. doi:10.1080/10615800310000112560

Cerin, E., Szabo, A., Hunt, N., & Williams, C. (2000). Temporal patterning of competitive emotions: A critical review. *Journal of Sports Sciences*, *18*(8), 605–626. doi:10.1080/02640410050082314

Chajut, E., & Algom, D. (2003). Selective attention improves under stress: Implications for theories of social cognition. *Journal of Personality and Social Psychology, 85*(2), 231–248.

Donders, F. C. (1969). On the speed of mental processes. *Acta Psychologica, 30*, 412–431.

Enander, A. E. (1989). Effects of thermal stress on human performance. Scandinavian *Journal of Work, Environment & Health, 15*, 27–33. Retrieved from http://www.jstor.org566699/stable/40965606

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance. *Emotion, 7*(2), 336–353. doi:10.1037/1528-3542.7.2.336

Friedland, N., Keinan, G., & Regev, Y. (1992). Controlling the uncontrollable: Effects of stress on illusory perceptions of controllability. *Journal of Personality and Social Psychology*, *63*(6), 923–931.

Friedman, A. (1979). Framing pictures: The role of knowledge in automatized encoding and memory for gist. *Journal of Experimental Psychology. General, 108*(3), 316–355.

Good, D. (2014). Predicting real-time adaptive performance in a dynamic decision-making context. *Journal of Management & Organization, 20*(6), 715–732. doi:10.1017/jmo.2014.54

Grant, E. R., & Spivey, M. J. (2003). Eye movements and problem solving: Guiding attention guides thought. *Psychological Science, 14*(5), 462–466.

Guilford, J. P., Christensen, P. R., Merrifield, P. R., & Wilson, R. C. (1978). *Alternate uses: Manual of instructions and interpretations*. Orange, CA: Sheridan Psychological Services.

Kahneman, D. (1973). Attention and effort. New Jersey: Prentice Hall. Retrieved from https://www.princeton.edu/~kahneman/docs/attention_and_effort/Attention_lo_quality.pdf

Kasarskis P., Stehwien J., Hickox J, Aretz A., & Wickens C. Comparison of expert and novice scan behaviors during VFR flight. The 11th International Symposium on Aviation Psychology, Columbus, OH. Association for Aviation Psychology; 2001. Retrieved rom https://www.researchgate.net/publication/243787467_Comparison_of_Expert_and_Novice_Scan_Behaviors_During_VFR_Flight

Kennedy, Q., Fricker, R., Appleget, J., Alt, J., Patton, S., Whitaker, L., Nesbitt, P., Yang, J. H., Huston, J. (2013). *Understanding optimal decision making in war-gaming*. TRADOC Analysis Center, Monterey, CA.

Kirk, R. E., & Hecht, E. (1963). Maintenance of vigilance by programmed noise. *Perceptual and Motor Skills, 16,* 553–560. doi:10.2466/pms.1963.16.2.553

Knoblich, G., Ohlsson, S., & Raney, G. E. (2001). An eye movement study of insight problem solving. *Memory & Cognition, 29*(7), 1000–1009.

Konstantopoulos, P., Chapman, P., & Crundall, D. (2010). Driver's visual attention as a function of driving experience and visibility. Using a driving simulator to explore drivers' eye movements in day, night and rain driving. *Accident; Analysis and Prevention, 42*(3), 827–834. doi:10.1016/j.aap.2009.09.022

Kundel, H. L., & La Follette, P. S. (1972). Visual search patterns and experience with radiological images. *Radiology, 103*(3), 523–528. doi:10.1148/103.3.523

Lavine, R. A., Sibert, J. L., Gokturk, M., & Dickens, B. (2002). Eye-tracking measures and human performance in a vigilance task. *Aviation, Space, and Environmental Medicine, 73*(4), 367–372.

Lieberman, H. R., Bathalon, G. P., Falco, C. M., Georgelis, J. H., Morgan, I. I. I., Niro, P., & Tharion, W. J. (2002). The "fog of war": Documenting cognitive decrements associated with the stress of combat. In *Proceedings of the 23rd Army Science Conference*. Orlando, FL. Retrieved from http://projects.mindtel.com/2005/SDSU.Geol600.Sensor_Networks/sensornets.refs/2003.%20ASC.%20Army%20Studies%20Conference/IO-01.pdf

Liu, Y., Hsueh, P., Lai, J., Sangin, M., Nüssli, M., & Dillenbourg, P. (2009). Who is the expert? Analyzing gaze data to predict expertise level in collaborative applications. Paper presented at the 2009 IEEE International Conference on Multimedia and Expo (ICME'09). IEEE Press, Piscataway, NJ. Retrieved from http://dl.acm.org/citation.cfm?id=1698924.1699144

Loftus, G. R., & Mackworth, N. H. (1978). Cognitive determinants of fixation location during picture viewing. Journal of Experimental Psychology. *Human Perception and Performance, 4*(4), 565–572.

MacLeod, C. M. (1991). Half a century of research on the stroop effect: An integrative review. *Psychological Bulletin, 109*(2), 163–203.

Maltz, M., & Shinar, D. (1999). Eye movements of younger and older drivers. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 41*(1), 15–25. doi:10.1518/001872099779577282

Marshall, S. P. (2007). Identifying cognitive state from eye metrics. *Aviation, Space, and Environmental Medicine, 78*(5 Suppl), 165.

Matthews, G., & Desmond, P. A. (2002). Task-induced fatigue states and simulated driving performance. *The Quarterly Journal of Experimental Psychology Section A, 55*(2), 659–686. doi:10.1080/02724980143000505

McCarley, J.S., & Kramer, A F. (2007). Eye movement as a window on perception and cognition. In R. Parasuraman, & and M. Rizzo, (Eds.), *Neuroergonomics: The brain at work* (pp. 95–112). New York: Oxford University Press

McPhee, L. C., Scialfa, C. T., Dennis, W. M., Ho, G., & Caird, J. K. (2004). Age differences in visual search for traffic signs during a simulated conversation. *Human Factors, 46*(4), 674–685.

Milham, M. P., Erickson, K. I., Banich, M. T., Kramer, A. F., Webb, A., Wszalek, T., & Cohen, N. J. (2002). Attentional control in the aging brain: Insights from an fMRI study of the Stroop task. *Brain and Cognition, 49*(3), 277–296.

Ozel, F. (2001). Time pressure and stress as a factor during emergency egress. *Safety Science, 38*(2), 95–107. doi:10.1016/S0925-7535(00)00061-8

Pieterse, G. (n.d.) What's stress got to do with it? [Blog post]. Retrieved from http://haleo.co.uk/wordpress/whats-stress-got-to-do-with-it/

Poole, A., Ball, L. J., & Phillips, P. (2005). In search of salience: A response-time and eye-movement analysis of bookmark recognition. In S. Fincher, P. Markopoulos, D. Moore, & R. Ruddle, (Eds.), *People and computers XVIII — Design for life* (pp. 363–378) Springer, London. Retrieved from http://link.springer.com/chapter/10.1007/1-84628-062-1_23

R Core Team. (2015). *R: A language and environment for statistical computing. R Foundation for Statistical Computing*. Vienna, Austria. URL http://www.R-project.org/.

Robert, J. & Hockey, G. (1997). Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. *Biological Psychology, 45*(1–3), 73–93. doi:10.1016/S0301-0511(96)05223-4

Sarter, N. B., Mumaw, R. J., & Wickens, C. D. (2007). Pilots' monitoring strategies and performance on automated flight decks: An empirical study combining behavioral and eye-tracking data. *Human Factors, 49*(3), 347–357.

Schriver, A. T., Morrow, D. G., Wickens, C. D., & Talleur, D. A. (2008). Expertise differences in attentional strategies related to pilot decision making. *Human Factors, 50*(6), 864–878.

Scialfa C.T., Jenkins L., Hamaluk E., Skaloud P. (2000). Aging and the development of automatic conjunction search. *Journal of Gerontology: Psychological Sciences, 55*(1), 27–46. doi:10.1093/geronb/55.1.P27

Shucard, J. L., McCabe, D. C., & Szymanski, H. (2008). An event-related potential study of attention deficits in posttraumatic stress disorder during auditory and visual go/NoGo continuous performance tasks. *Biological Psychology, 79*(2), 223–233. doi:10.1016/j.biopsycho.2008.05.005

Silvia, P. J., Winterstein, B. P., Willse, J. T., Barona, C. M., Cram, J. T., Hess, K. I., . . . Richard, C. A. (2008). Assessing creativity with divergent thinking tasks: Exploring the reliability and validity of new subjective scoring methods. *Psychology of Aesthetics, Creativity, and the Arts, 2*(2), 68–85. doi:10.1037/1931-3896.2.2.68

Simmonds, D. J., Pekar, J. J., & Mostofsky, S. H. (2008). Meta-analysis of go/no-go tasks demonstrating that fMRI activation associated with response inhibition is task-dependent. *Neuropsychologia, 46*(1), 224–232. doi:10.1016/j.neuropsychologia.2007.07.015

Spivey, M.J., Richardson, D.C., & Fitneva, S.A. (2004). Thinking outside the brain: Spatial indices to visual and linguistic information. *The Interface of Language, Vision and Action: Eye Movements and the Visual World*, (pp.161-189). New York: Psychology Press.

Staal, M. A. (2004). Stress, cognition, and human performance: A literature review and conceptual framework. (NASA/TM -2004-212824). Retrieved from http://ntrs.nasa.gov/search.jsp?R=20060017835

Steinhauser, M., Maier, M., & Hübner, R. (2007). Cognitive control under stress : How stress affects strategies of task-set reconfiguration. Retrieved from http://nbn-resolving.de/urn:nbn:de:bsz:352-opus-58745

Stokes, A. F., & Kite, K. (2001). On grasping a nettle and becoming emotional. In P.A. Hancock & P.A. Desmond (Eds.), *Stress, workload and fatigue*. Mahwah, NJ: L. Erlbaum

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*(6), 643–662. doi:10.1037/h0054651

Sullivan, J., Yang, J. H., Day, M., & Kennedy, Q. (2011). Training simulation for helicopter navigation by characterizing visual scan patterns. *Aviation, Space, and Environmental Medicine, 82*(9), 871–878. doi:10.3357/ASEM.2947.2011

Tipper, S. P. (1985). The negative priming effect: Inhibitory priming by ignored objects. *The Quarterly Journal of Experimental Psychology, 37* (4), 571–90. doi: 10.1080/14640748508400920

United States Army Combined Arms Center. (2014). *The human dimension white paper: a framework for optimizing human performance*. Retrieved from http://usacac.army.mil/pubs/Force-2025-and-Beyond-Human-Dimension

Van Galen, G. P., & van Huygevoort, M. (2000). Error, stress and the role of neuromotor noise in space oriented behaviour. *Biological Psychology, 51*(2-3), 151–171.

Vine, S. J., Uiga, L., Lavric, A., Moore, L. J., Tsaneva-Atanasova, K., & Wilson, M. R. (2014). Individual reactions to stress predict performance during a critical aviation incident. *Anxiety, Stress, & Coping*, *28*(4), 467–477. doi:10.1080/10615806.2014.986722

Virk, S. A., & Eslick, G. D. (2015). Brief report. *Epidemiology, 26*(5), 769–773. doi:10.1097/EDE.0000000000000326

Wickham, H. (2009). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.

Wikman, A. & Summala, H. (2005). Aging and time-sharing in highway driving. *Optometry and Vision Science, 82*(8). Retrieved from https://trid.trb.org/view.aspx?id=795152

Wilson, M. R. (2012). Anxiety: Attention, the brain and body performance. In S. Murphy (Ed.), *The Oxford Handbook of Sport and Performance Psychology* (pp. 173–190). New York, NY: Oxford University Press.

Wofford, J. C. (2001). Cognitive-affective stress response: Effects of individual stress propensity on physiological and psychological indicators of strain. *Psychological Reports*, *88*(3 Pt 1), 768–784. doi:10.2466/pr0.2001.88.3.768

Wofford, J. C., Goodwin, V. L., & Daly, P. S. (1999). Cognitive-affective stress propensity: A field study. *Journal of Organizational Behavior, 20*(5), 687–707. doi:AID-JOB926>3.0.CO;2-5

Wofford, J. C., & Goodwin, V. L. (2002). The linkages of cognitive processes, stress propensity, affect, and strain: Experimental test of a cognitive model of the stress response. *Personality and Individual Differences*, *8*(32), 1413–1430. Retrieved from https://www.infona.pl//resource/bwmeta1.element.elsevier-b162ca59-6b17-3e46-b7e2-64c17a27df05

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, *18*(5), 459–482. doi:10.1002/cne.920180503

Zelinsky, G. J. (2008). A theory of eye movements during target acquisition. *Psychological Review, 115*(4), 787–835. doi:10.1037/a0013118

INITIAL DISTRIBUTION LIST

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California